

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 16:04:46 ; Search time 41 Seconds

(without alignments)  
96.784 Million cell updates/sec

Title: US-09-774-940A-1

Sequence: 1 YQCEICGKSPDKSNLTRRLHTG 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq\_19Jun03:\*  
1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	77.7	29	20	AAV33363 Consensus zinc fin
2	108	77.7	29	20	AAW78386 Synthetic Cys2-His
3	108	77.7	29	20	AAW84303 Consensus zinc fin
4	108	77.7	29	20	AAW87699 Consensus zinc fin
5	108	77.7	29	21	AAV94377 Consensus zinc fin
6	108	77.7	29	22	AAU05197 Zinc finger consen
7	108	77.7	29	22	AAW63327 Zinc finger protei
8	108	77.7	29	22	AAW63327 Amino acid sequenc
9	106	76.3	242	24	ABU19775 Zinc finger consen Human MP21 protein

10	106	76.3	448	22	AAW60249 Human protein SEQ
11	106	76.3	480	23	AAW48967 Human zinc finger
12	106	76.3	555	22	AAW79265 Human protein SEQ
13	106	76.3	555	22	AAW93576 Human protein sequ
14	105	75.5	593	22	ABW70360 Drosophila melanog
15	104	74.8	362	23	ABW70266 Human prey protein
16	104	74.8	525	22	ABW47725 Human liver protein
17	104	74.8	525	22	ABW32702 Human peptide #353
18	104	74.8	525	22	ABW32702 Human peptide #353
19	104	74.8	525	22	ABW32702 Human peptide #353
20	104	74.8	525	22	ABW32702 Human peptide #353
21	104	74.8	525	22	ABW32702 Human peptide #353
22	104	74.8	525	22	ABW32702 Human peptide #353
23	104	74.8	525	22	ABW32702 Human peptide #353
24	104	74.8	525	22	ABW32702 Human peptide #353
25	104	74.8	525	22	ABW32702 Human peptide #353
26	104	74.8	525	22	ABW32702 Human peptide #353
27	102	73.4	25	22	ABW32702 Human peptide #353
28	102	73.4	25	22	ABW32702 Human peptide #353
29	102	73.4	25	22	ABW32702 Human peptide #353
30	102	73.4	25	22	ABW32702 Human peptide #353
31	102	73.4	25	22	ABW32702 Human peptide #353
32	102	73.4	25	22	ABW32702 Human peptide #353
33	102	73.4	25	22	ABW32702 Human peptide #353
34	102	73.4	25	22	ABW32702 Human peptide #353
35	102	73.4	25	22	ABW32702 Human peptide #353
36	102	73.4	25	22	ABW32702 Human peptide #353
37	101	72.7	208	24	ABW1413 Human DITP zinc f
38	101	72.7	338	22	AAW93945 Human polypeptide,
39	101	72.7	517	22	ABW33375 Human human diagno
40	101	72.7	577	23	AAW14680 Human transcriptio
41	100	71.9	26	20	AAW87559 Consensus zinc fin
42	100	71.9	236	23	ABW42018 Human ovarian anti
43	100	71.9	1050	22	ABW09685 Novel human diagno
44	100	71.9	1212	22	ABW00399 Novel human diagno
45	100	71.9	1230	22	AAW30831 Novel human secret

#### ALIGNMENTS

RESULT 1  
AAV33363  
ID AAV33363 standard; peptide; 29 AA.  
XX  
AC AAV33363;  
XX  
DT 01-DEC-1999 (first entry)  
XX  
DE Consensus zinc finger peptide motif 2.  
XX  
KW Zinc finger; DNA binding; Cys2-His2 class; 5-methylcytosine; meC;  
XX  
XX diagnostic; detection; chimera.  
OS Unidentified.  
XX  
XX W09947656-A2.  
XX  
PD 23-SBP-1999.  
XX  
PP 17-MAR-1999; 99WC-GB00816.  
XX  
PR 17-MAR-1998; 98GB-0005576.  
PR 31-MAR-1998; 98GB-0006895.  
PR 03-APR-1998; 98GB-0007246.  
PA (MED1-) MEDICAL RES COUNCIL.  
PI Choc Y, Isalan M;  
XX  
XX WPI; 1999-562106/47.  
DR  
XX  
PT New zinc finger polypeptides that bind DNA containing modified bases,



Query Match 77.7%; Score 108; DB 20; Length 29;  
Best Local Similarity 76.0%; Pred. No. 1.1e-08;  
Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0

```
QY      1 YQCEICGKSFSDKSNLTRHLRIHTG 25
      2  |||||:||||| |||||
Db      2 YKCECGKAFSQKSNLTRHQRIHTG 26
```

RESULT 4	
AAW87699	
ID	AAW87699 standard; peptide; 29 AA
XX	

DT 09-MAR-1999 (first entry)  
 VV

Consensus zinc finger sequence of the Cys2-His2 zinc finger class

LM zinc finger; Cys2-His2 zinc finger class; chimeric restriction enzyme  
 MN fusion protein; gene therapy; nucleic acid binding protein.  
 NY

OS Synthetic.

PN W09853057-A1.

PD 26-NOV-1998

PF 26-MAY-1998; 98WO-GB01510.  
XX

PR 23-MAY-1997; 97GB-0010809.  
XX

PA (MED-) MEDICAL RES COUNCIL.  
XX

PL Choo Y, Isalan M, Klug A, XX

DR WPL: 1999-024577/02.  
XX

PT new library of nucleic acid binding zinc finger polypeptide(s)  
PT each polypeptide comprising more than one zinc finger which is  
PT partially randomised, useful for detecting a target nucleic acid  
PT sequence

PS Claim 20; Page 45; 56pp; English

The present sequence represents a consensus zinc finger sequence of the Cys2-His2 zinc finger class. It is used to create a zinc finger polypeptide library in which each polypeptide comprises more than one zinc finger which has been at least partially randomized. Zinc finger proteins bind to particular nucleic acid targets. The proteins can be used for determining the presence of a target nucleic acid. The proteins of the invention can be used in the manufacture of chimeric restriction enzymes, in which a nucleic acid cleaving domain is fused to a nucleic acid binding domain comprising a zinc finger. Fusion proteins comprising a binding protein and an integrase, e.g. viral integrase, can be used to target nucleic acid sequences *in vivo*. In gene therapy applications, the method may be targeted to the delivery of functional genes into defective genes, or the delivery of nonsense nucleic acid. Genes may also be delivered to known repetitive stretches of nucleic acid e.g. centromeres, together with an activating sequence such as an ICR. Nucleic acid binding proteins can be specifically used to knock-out cells having mutant proteins, e.g. mutant ras. They can also be used to modulate the action of transcription factors, e.g. the activity of HIV tat may be reduced by binding proteins specific for HIV TAR. The new binding proteins may also be coupled to toxic molecules, e.g. nucleases, which are capable of selectively destroying cells which comprise a mutation in their endogenous nucleic acid. The products can be used in the treatment of infections.

Sequence	29	AA;
Query Match	77.7%	Score 108; DB 20; Length 29

Best Local Similarity 76.0%; Pred. No. 1.1e-08;  
Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

QY      1 YQCEICGKSFSDKSNLTRHLRIHTG 25
      2 1: |||: ||| ||| ||| |||
Db      2 YKCEGKAFSQKSNLTRHQRIHTG 26

```

RESULT 5	
AAy94377	
ID	AAy94377 standard; Peptide; 29 AA
XY	

DT 22-AUG-2000 (First entry)  
 VV

Zinc finger consensus structure #2.

**KW** Zinc finger; DNA binding protein; polysome display;  
**KW** binding protein selection.  
**XY**

OS Unidentified

PN WO200027878-A1.

PD 18-MAY-2000.

PF 09-NOV-1999; 99MO-GB03730.  
XX

FR 09-NOV-1998; 98GB-0024544;  
XX

(GEND-) GENDAQ L1'D

XX  
F1  
C100 1;

DR  
XX  
NET! 2000-3/6494/32

PT A polysome display-based technique for producing and selecting zinc  
PT finger nucleic acid binding proteins with desired binding  
PT characteristics -

PS Disclosure; Page 10; 49pp; English.  
xy

The present sequence is a zinc finger consensus sequence which may be incorporated into a zinc finger transcription unit comprising a T7 promoter, ribosome binding site, zinc finger coding sequence and a linker/stalling sequence. By modifying the coding sequence, the zinc finger protein may be varied at one or more positions and variants which bind to a target nucleic acid sequence may be selected by polyclonal display. In this way, zinc finger binding proteins with desired binding characteristics can be isolated without resorting to phage display techniques. The present sequence may be prepared by comparing the sequences of known zinc fingers, irrespective of whether their binding domain is known.

**SQ Sequence 29 AA;**

Query Match	77.7%;	Score 108;	DB 21;	Length 29;
Best Local Similarity	76.0%;	Pred. No. 1.1e-08;		
Matches 19; Conservative	2;	Mismatches 4;	Indels 0;	Gaps 0

Dy YQCEICGKSFSDKSNTLRLRHITG 25  
Db YKCSECGKAFSQKSNLTRHRIHTG 26

RESULT 6  
AAU05197  
ID AAU05197 standard; Protein; 29 AA  
rv

DT 24-OCT-2001 (first entry)

Zinc finger protein consensus structure #2.

TFFIIA/Zif-VI6; sub-cellular trafficking; signal transduction; diagnostic; gene therapy; agricultural technology; zinc finger. Synthetic.

MO200153479-A2.

26-JUL-2001.

18-JAN-2001; 2001WO-GB00187.

24-JAN-2000; 2000GB-0001578.  
PR 24-JAN-2000; 2000GB-0001582.  
PR 30-MAY-2000; 2000KO-GB02071.  
PR 30-MAY-2000; 2000WC-GB02080.  
FR 07-DEC-2000; 2000GC-0029901.

(GEND-) GENDAQ LTD.

Choo Y, Ullman CG, Moore M;  
WPI: 2001-457607/49.

Selecting switching system components comprising first molecule, second molecule and ligand, involves determining degree of binding between the molecules in the presence and absence of the ligand -

Disclosure; Page 54; 19pp; English.

The invention relates to a method of selecting components of a switching system comprising first molecule (M1), selecting (i) a ligand (L2) which is capable of modulating the interaction between a nucleic acid binding molecule (NBM) and a target nucleic acid (TN). NBM, TN and the ligands are useful for regulating a biological process selected from sub-cellular trafficking and signal transduction, and for regulating transcription or translation from a nucleic acid sequence comprising TN to which NBM binds in a manner modulatable by the ligand. They are also useful for modulating a biological process affecting one or more genes, such as in transcription, translation, phosphorylation, replication, transport, splicing, integration and recombination, in a host cell. The polypeptides or ligands selected from (S) are useful for regulating a biological process involving binding of first polypeptide with second polypeptide, selected from sub-cellular trafficking and signal transduction. (i) is also useful for screening any or all the components of the gene switch system or protein switch system, and for modulating the interaction between the components of the switching system. The binding molecules are useful as diagnostic and research tools, and for switching or modulating gene expression, especially in gene therapy applications and agricultural technology applications. The ligands are also useful in the treatment of diseases. Multiple target nucleic acids or polypeptides could be used in a single selective step, thus enabling multiple nucleic acid or polypeptide binding molecules to be isolated simultaneously, even in the same physical vessel. The present sequence represents the amino acid sequence of zinc finger protein consensus structure #2 used in the method of the invention.

Query Match 77.7%; Score 108; DB 22; Length 29;  
Best Local Similarity 76.0%; Pred. No. 1,1e-08;  
Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0

1 YCCETCGKSFSDKSNTTRRLRIHTG 25  
| : | | | | | | | | |  
2 YKCEGCKAFSSCKSNLTRQRIRHTG 26

RESULT 7  
AAAG63227 standard; peptide; 29 AA

XX	AC	AA63227;
XX	DT	01-OCT-2001 (first entry)
XX	DE	Amino acid sequence of a consensus zinc finger peptide structure.
XX	KW	Transcription; plant cell; zinc finger polypeptide; pest resistance;
XX	OS	Synthetic.
XX	RN	WO2001S3478-A2.
XX	PD	26-JUL-2001.
XX	PF	22-JAN-2001; 2001WO-US02051.
XX	PR	24-JAN-2000; 2000GB-0000001.
XX	PR	07-DEC-2000; 2000US-0000001.
XX	PA	(GEND-) GENDAQ LTD.
XX	PA	(UYRQ) UNIV ROCKEFELLER.
XX	PA	(CHOO/) CHOO Y.
XX	PA	(ULIM/) ULLMAN C G.
XX	PA	(CHUA/) CHUA N.
XX	PA	(SANC/) SANCHEZ J P.
XX	PI	Choo Y, Ullman CG, Chua N, Sanchez JP,
XX	DR	WPI; 2001-476119/51.
XX	PT	Regulating transcription in a plant cell, useful for regulating gene
XX	PT	expression in a plant, particularly for improving e.g., resistance to
XX	PT	pest, disease or herbicide, by introducing an engineered zinc finger
XX	PT	polypeptide into the cell -
XX	PS	Disclosure, Page 13; 62pp; English.
XX	CC	The specification describes a method for regulating transcription in
XX	CC	a plant cell from a DNA sequence having a target DNA operably linked
XX	CC	to a coding sequence. The method comprises introducing an engineered
XX	CC	zinc finger polypeptide into the plant cell. The polypeptide binds to
XX	CC	the target DNA and modulates transcription of the coding sequence. The
XX	CC	method is useful for regulating gene expression in a plant. The method
XX	CC	is particularly useful in biotechnology, particularly for improving
XX	CC	various properties of plants, e.g. resistance to pests, diseases or
XX	CC	herbicides, and for improving various seed and fruit traits. The present
XX	CC	sequence represents a consensus zinc finger peptide structure.
XX	SQ	Sequence 29 AA;
XX	Query Match	77.7%; Score 108; DB 22; Length 29;
XX	Best Local Similarity	76.0%; Fred. No. 1.1e-08;
XX	Matches 19; Conservative	2; Mismatches 4; Indels 0; Gaps 0.
OY	1	YOCFICGKSPSDKSNLTRHLRIHTG 25   :                           Db 2 YKCSECKAFSGKSNLTRHGRIRHTG 26
XX	RESULT 8	
XX	AAAB20022	
XX	ID	AAAB20022 standard; Peptide; 29 AA.
XX	AAAB20022;	
XX	23-APR-2001	(first entry)
XX	DB	Zinc finger consensus structure.
XX	KW	Molecular switch; gene switch; protein switch; zinc finger.

OS Synthetic.  
 XX WO200100815-A1.  
 PN  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 30-MAY-2000; 2000MO-GB02080.  
 XX  
 PR 28-MAY-1999; 99GB-0012635.  
 PR 24-JAN-2000; 2000GB-0001582.  
 XX  
 PA (GEND-) GENDAQ LTD.  
 XX  
 PI Choo Y, Ullman CGU;  
 DR WPI; 2001-123005/13.  
 XX  
 PT Selecting gene or protein switches for use in modulating gene  
 PT expression, comprises selecting complexes of two molecules and a ligand  
 PT whose binding differs in the presence or absence of the ligand  
 XX  
 XX Disclosure; Page 29; 125pp; English.  
 XX  
 CC The present sequence is that of a consensus zinc finger structure.  
 CC The invention relates to molecular switches and methods for  
 CC identifying and selecting such switches for use in modulating gene  
 CC expression e.g. in gene therapy. Gene switches use molecules  
 CC capable of binding a specific DNA sequence in a ligand-dependent  
 CC manner. In protein switches, 2 protein binding partners bind in a  
 CC molecule, which is modifiable by a ligand. Polypeptide binding  
 CC molecules, such as zinc fingers may be employed in the regulation  
 CC of gene transcription, e.g. by specific cleavage of nucleic acid  
 CC sequences using a fusion protein comprising a zinc finger-targeting  
 CC domain and a nucleic acid cleavage domain, or by fusion of an  
 CC transcriptional effector domain to a zinc finger, to activate or  
 CC repress transcription.  
 XX  
 SQ Sequence 29 AA;  
 Query Match 77.7%; Score 108; DB 22; Length 29;  
 Best Local Similarity 76.0%; Pred. No. 1.1e-08;  
 Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 YQCEICGKSPDKSNLTRHRIHTG 25  
 DB 2 YKCEGCKAFSGKSNLTRHRIHTG 26  
 RESULT 9  
 ID ABJ19775 standard; Protein; 242 AA.  
 XX  
 AC ABJ19775;  
 XX  
 DT 03-APR-2003 (first entry)  
 XX  
 DE Human MP21 protein Kox21 SEQ ID NO 55.  
 XX  
 KM Cytostatic; p21 pathway modulating agent; cancer; angiogenic; apoptotic;  
 KM cell proliferation disorder; MP21.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003006990-A1.  
 PD 23-JAN-2003.  
 XX  
 PF 10-JUL-2002; 2002WO-US21549.  
 XX  
 PR 12-JUL-2001; 2001US-305017P.  
 PR 10-OCT-2001; 2001US-328491P.  
 PR 15-FEB-2002; 2002US-357452P.  
 XX

PA (EXEL-) EXELIXIS INC.  
 XX  
 PI Friedman L, Plozman GD, Belvin M, Li D, Funke RP;  
 XX  
 DR WPI; 2003-221779/21.  
 DR N-PSDB; ABT17049.  
 XX  
 PT Identifying candidate p21 pathway modulator, by contacting an assay  
 PT system having modifiers of p21 polypeptide or gene with a test agent to  
 PT provide a reference activity in system and detecting test agent-biased  
 PT activity -  
 XX  
 XX Examples; Page 176-177; 199pp; English.  
 XX  
 CC The invention relates to a novel method for identifying a candidate p21  
 CC pathway modulating agent. The novel method comprises contacting an assay  
 CC system, comprising a purified MP21 polypeptide (modifier of p21) or  
 CC nucleic acid, with a test agent under conditions, so that but for the  
 CC presence of a test agent, the assay system provides a reference activity  
 CC and detection of test agent-biased activity of the assay system. The  
 CC novel method of the invention is useful for identifying a candidate p21  
 CC pathway modulating agent. The invention also includes a method for  
 CC modulating the p21 pathway of a cell, and a method for diagnosing a  
 CC disease e.g. cancer in a patient. The identified modulators are useful in  
 CC diagnosis, therapy and pharmaceutical development. The modulators are  
 CC useful in a variety of diagnostic and therapeutic applications including  
 CC angiogenic, apoptotic and cell proliferation disorders. This sequence  
 CC represents an MP21 protein of the invention.  
 XX  
 SQ Sequence 242 AA;  
 Query Match 76.3%; Score 106; DB 24; Length 242;  
 Best Local Similarity 72.0%; Pred. No. 2.1e-07;  
 Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 YQCEICGKSPDKSNLTRHRIHTG 25  
 DB 92 YECNEGCKSRSEKSTLTKHRIHTG 116  
 RESULT 10  
 ID AAM80249 standard; Protein; 448 AA.  
 XX  
 AC AAM80249;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human protein SEQ ID NO 3895.  
 XX  
 KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KM nervous system disorder; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157190-A2.  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,  
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW,  
PI Xue AU, Yang Y, Wejhrman T, Goodrich R,  
XX WPI: 2001-476283/51.  
DR N-PSDB; AAK53382.  
XX  
PT Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
XX  
PS Claim 20; Page 463; 6221pp; English.  
XX  
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
CC (AAM80020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.  
XX  
SQ Sequence 448 AA;  
XX  
Query Match 76.3%; Score 106; DB 22; Length 448;  
Best Local Similarity 68.0%; Pred. No. 4.2e-07;  
Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
OY 1 YQCEIGKSPSFKSNLTRHRIHTG 25  
DB 339 YKCDLGGKVFQKSNLRAHWRVHTG 363  
XX  
RESULT 11  
AAM48967  
ID AAM48967 standard; Protein: 480 AA.  
XX  
AC AAM48967;  
XX  
DT 25-APR-2002 (first entry)  
XX  
XX Human zinc finger protein 53.  
XX  
XX Human zinc finger protein 53; cancer; nervous system disease;  
XX development disorder; metabolic disease; inflammation; haemopathy;  
XX immunological disease; HIV infection; gene therapy.  
XX  
XX Homo sapiens.  
XX  
XX CN1314368-A.  
XX  
XX 26-SEP-2001.  
XX  
XX 17-MAR-2000; 2000CN-0114979.  
XX  
XX 17-MAR-2000; 2000CN-0114979.  
XX  
XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.  
XX  
XX Mao Y, Xie Y;  
XX  
XX WPI: 2002-056224/08.  
XX  
XX N-PSDB; AAL41725.  
XX  
XX New polypeptide-human zinc finger protein 53 and polynucleotide for  
XX coding such polypeptide -  
XX

PS Claim 1; Page 25-26(Disclosure); 33pp; Chinese.  
XX  
XX The present invention provides the protein and coding sequences of human  
CC zinc finger protein 53. The sequences can be used in the treatment of  
CC cancer, haemopathy, nervous system disorders, development disorders,  
CC metabolic disorders, inflammation, immunological diseases and HIV  
CC infection. The present sequence is the protein of the invention.  
XX  
SQ Sequence 480 AA;  
XX  
Query Match 76.3%; Score 106; DB 23; Length 480;  
Best Local Similarity 68.0%; Pred. No. 4.5e-07;  
Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
OY 1 YQCEIGKSPSFKSNLTRHRIHTG 25  
DB 169 YKCDLGGKVFQKSNLRAHWRVHTG 193  
XX  
RESULT 12  
AAM79265  
ID AAM79265 standard; Protein: 555 AA.  
XX  
AC AAM79265;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
XX Human protein SEQ ID NO 1927.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
XX tissue growth factor; immunomodulatory; cancer; leukaemia;  
XX nervous system disorder; arthritis; inflammation.  
XX  
XX Homo sapiens.  
XX  
XX WO200157190-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 05-FEB-2001; 2001WO-US04098.  
XX  
XX 03-FEB-2000; 2000US-0496914.  
XX  
XX 27-APR-2000; 2000US-0560875.  
XX  
XX 20-JUN-2000; 2000US-0598075.  
XX  
XX 19-JUL-2000; 2000US-0620325.  
XX  
XX 01-SEP-2000; 2000US-0654936.  
XX  
XX 15-SEP-2000; 2000US-0653561.  
XX  
XX 20-OCT-2000; 2000US-0653325.  
XX  
XX 30-NOV-2000; 2000US-0728422.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX  
XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,  
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW,  
PI Xue AU, Yang Y, Wejhrman T, Goodrich R,  
XX WPI: 2001-476283/51.  
XX  
XX N-PSDB; AAK52398.  
XX  
XX Nucleic acids encoding polypeptides with cytokine-like activities,  
XX useful in diagnosis and gene therapy -  
XX  
PS Claim 20; Page 4329-4330; 6221pp; English.  
XX  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 CC  
 SQ Sequence 555 AA;  
 Query Match 76.3%; Score 106; DB 22; Length 555;  
 Best Local Similarity 68.0%; Pred. No. 5.3e-07;  
 Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 YQCEICGKSFSDKSNLTTRHLRIHTG 25  
 244 YKCDLCGKVFQKSNLARHWRVHTG 268  
 RESULT 13  
 AAB93576  
 ID AAB93576 standard; Protein; 555 AA.  
 XX  
 AC AAB93576;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO:12988.  
 XX  
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 12988; 2537pp + CD ROM; English.  
 CC  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises at least 15 nucleotides and the combination  
 CC of the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 CC  
 SQ Sequence 555 AA;  
 Query Match 76.3%; Score 106; DB 22; Length 555;  
 Best Local Similarity 68.0%; Pred. No. 5.3e-07;  
 Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 YQCEICGKSFSDKSNLTTRHLRIHTG 25  
 244 YKCDLCGKVFQKSNLARHWRVHTG 268  
 RESULT 14  
 ABB70360  
 ID ABB70360 standard; Protein; 593 AA.  
 XX  
 AC ABB70360;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 37872.  
 XX  
 KM Drosophila; developmental biology; cell signalling; insecticide;  
 KM pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN W0200171042-A2.  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 XX  
 DR N-PSDB; ABL14463.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 37872; 21pp + Sequence Listing; English.  
 CC  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABU30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (AAB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pot\_sequences.  
 XX  
 SQ Sequence 593 AA;  
 Query Match 75.5%; Score 105; DB 22; Length 593;  
 Best Local Similarity 64.0%; Pred. No. 8e-07;  
 Matches 16; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 1 YQCEICGKSFSDKSNLTRHLRIHTG 25  
 Db 445 YQCEICGKSFADRSNMTLHRLHSHG 469

## RESULT 15

ABG70216  
 ID ABG70216 standard, Protein, 362 AA.

XX ABG70216;

XX 21-OCT-2002 (first entry)

XX Human prey protein for Shigella ospG #2.

XX Prey protein; ospB; ospD1; ipad; ipaH9.8; ospG;

KM ospC1; Shigella; shigellosis; bacillary dysentery; antibacterial;  
 KM yeast two-hybrid system; protein-protein interaction; SID;  
 KM selected interacting domain; human.

XX Homo sapiens.

XX W0200257303-A2.

XX 25-JUL-2002.

XX 11-JAN-2002; 2002MO-EP00777.

XX 12-JAN-2001; 2001US-261130P.

XX (HYBR-) HYBRIGENICS.

XX Legrain P;

XX WPI: 2002-5599706/64.

XX N-PSDB; AB551609.

XX New complex of protein-protein interactions between a bait Shigella  
 PT flexneri polypeptide and a prey mammalian or human placenta polypeptide  
 PT for treating or preventing bacillary dysentery in a mammal or human -  
 XX

PS Claim 7; Page 137; 162pp; English.

XX The invention relates to a complex of protein-protein interactions  
 CC between a Shigella flexneri polypeptide (e.g. ospB, ospD1, ipad, ipaC,  
 CC ipaH9.8, ospG and ospC1) and a mammalian polypeptide defined in the  
 CC specification. The complexes are formed using the yeast two-hybrid  
 CC system. Also included are (1) a recombinant host cell expressing the  
 CC interactions between the Shigella flexneri polypeptide and a mammalian  
 CC polypeptide defined in the specification; (2) selecting a modulating  
 CC compound that inhibits or activates the protein-protein interactions;  
 CC (3) a modulating compound obtained from the method of (2); (4) a SID  
 CC (selected interacting domain) polypeptide or its fragment or variant  
 CC comprising the human polypeptides appearing as ABG70042-ABG70242;  
 CC (5) a SID polynucleotide or its fragment or variant comprising  
 CC encoding the above polypeptides a vector comprising (5);  
 CC (6) a recombinant host cell containing the vector; and (10) a protein  
 CC chip comprising Shigella flexneri polypeptide and a mammalian polypeptide  
 CC defined in the specification. A pharmaceutical composition comprising the  
 CC compound, polypeptide or polynucleotide is useful for treating or  
 CC preventing shigellosis (bacillary dysentery) in a human or mammal.  
 CC The present sequence represents a human prey protein isolated by the  
 CC yeast two-hybrid assay, forming a complex of the invention with a  
 CC shigella protein.  
 XX

SQ Sequence 362 AA;

Query Match 74.8%; Score 104; DB 23; Length 362;

Best Local Similarity 76.0%; Pred. No. 6.5e-07;

Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 YQCEICGKSFSDKSNLTRHLRIHTG 25

Db 108 YQCEICGKSFQDLNLTFRHRIHTG 132

Search completed: January 6, 2004, 16:10:40  
 Job time : 43 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 16:08:31 ; Search time 36 Seconds

(without alignments)  
179.203 Million cell updates/sec

Title: US-09-774-940A-1

Perfect score: 139  
Sequence: 1 YGCICGKSFSDKSNLTRHLRHNG 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	76.3	555	4 Q9NR2	Q9NR2 homo sapien
2	106	76.3	561	4 Q96B8	Q96B8 homo sapien
3	105	75.5	587	5 Q9NK9	Q9NK9 drosophila
4	105	75.5	608	11 Q8V08	Q8V08 mus musculu
5	105	75.5	608	11 Q8BK8	Q8BK8 mus musculu
6	105	75.5	608	11 Q8BK8	Q8BK8 mus musculu
7	104	74.8	576	11 Q8RQ3	Q8RQ3 mus musculu
8	104	74.8	576	4 Q9N22	Q9N22 mus musculu
9	104	74.8	606	4 Q9UJ4	Q9UJ4 homo sapien
10	104	74.1	475	4 Q8NA2	Q8NA2 homo sapien
11	102	73.4	106	4 Q9UFI	Q9UFI homo sapien
12	102	73.4	337	11 Q8R0X	Q8R0X mus musculu
13	102	73.4	496	11 Q8C53	Q8C53 mus musculu
14	102	73.4	496	11 Q8B55	Q8B55 mus musculu
15	102	73.4	530	4 Q8NA6	Q8NA6 homo sapien
16	102	73.4	686	4 Q9NR4	Q9NR4 homo sapien

17	102	73.4	824	11 Q921B4	Q921B4 mus musculu
18	102	73.4	874	11 Q9R163	Q9R163 mus musculu
19	102	73.4	913	4 Q9HC27	Q9HC27 homo sapien
20	101	72.7	338	4 Q8N79	Q8N79 homo sapien
21	101	72.7	404	4 Q8N7Q3	Q8N7Q3 homo sapien
22	101	72.7	485	4 Q8IY56	Q8IY56 homo sapien
23	101	72.7	530	4 Q8NEM1	Q8NEM1 homo sapien
24	101	72.7	576	4 Q8IWC4	Q8IWC4 homo sapien
25	101	72.7	595	4 Q8TF32	Q8TF32 homo sapien
26	100	71.9	282	4 Q9B8G1	Q9B8G1 homo sapien
27	100	71.9	282	4 Q9B8G1	Q9B8G1 homo sapien
28	100	71.9	335	4 Q9UMC5	Q9UMC5 homo sapien
29	100	71.9	475	11 Q9C286	Q9C286 mus musculu
30	100	71.9	637	4 Q9E1R2	Q9E1R2 mus musculu
31	99	71.2	386	11 Q8VC11	Q8VC11 mus musculu
32	99	71.2	391	4 Q8N2S2	Q8N2S2 homo sapien
33	99	71.2	393	4 Q75260	Q75260 homo sapien
34	99	71.2	418	4 Q9B9W8	Q9B9W8 homo sapien
35	99	71.2	477	4 Q9Y4B2	Q9Y4B2 homo sapien
36	99	71.2	626	4 Q8N3V1	Q8N3V1 homo sapien
37	99	71.2	626	4 Q9BX82	Q9BX82 homo sapien
38	98	70.5	208	4 Q8N4Z8	Q8N4Z8 homo sapien
39	98	70.5	229	4 Q8N506	Q8N506 homo sapien
40	98	70.5	230	4 Q9BPG0	Q9BPG0 homo sapien
41	98	70.5	446	4 Q9B8E5	Q9B8E5 homo sapien
42	98	70.5	454	4 Q9B8Q8	Q9B8Q8 homo sapien
43	98	70.5	461	11 Q8BIQ2	Q8BIQ2 mus musculu
44	98	70.5	493	4 Q60304	Q60304 homo sapien
45	98	70.5	555	11 Q921V2	Q921V2 mus musculu

## ALIGNMENTS

### RESULT 1

Q9NR2 PRELIMINARY; PRT; 555 AA.

AC Q9NR2;  
DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
DE Hypothetical protein FLJ1191.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M.,  
RA Watanabe S., Kimura K., Chida Y., Kodaira H., Kondo H., Murakami K.,  
RA Takahashi M., Chida Y., Ishida S., Murakami K., Ono Y., Takaguchi S.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuhara Y.,  
RA Ninomiya K., Iwayanagi T.,  
RT "NEO human cDNA sequencing project."  
RT Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AK002053; BAA92059.1; ..  
DR HSSP: P08045; 1ZNF.  
DR InterPro: IPR001909; KRAB.  
DR InterPro: IPR007087; Znf\_C2H2.  
DR Pfam: PF01352; KRAB; 1.  
DR Pfam: PF00096; zf-C2H2; 11.  
DR ProDom: PD000003; Znf\_C2H2; 11.  
DR SMART: SM00349; KRAB\_1.  
DR SMART: SM00355; ZNF\_C2H2; 11.  
DR PROSITE: PS00805; KRAB; 1.  
DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 11.  
DR PROSITE: PS00157; ZINC\_FINGER\_C2H2\_2; 11.  
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.  
SQ SEQUENCE 555 AA; 64101 MW; D3B5B92BBEB2604 CRC64;

Query Match 76.3%; Score 106; DB 4; Length 555;  
 Best Local Similarity 68.0%; Pred. No. 6.2e-08;  
 Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 YQCEICGKSPDSKNTLRHRIHTG 25  
 Db 244 YKCDLCKGKFSQSKSNLAKHMKVHTG 268

## RESULT 2

Q96B88 PRELIMINARY; PRT; 561 AA.

DT 01-DEC-2001 (TRENBLREL. 15, Created)

DT 01-DEC-2001 (TRENBLREL. 15, Last sequence update)

DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)

DB Hypothetical protein (Zinc finger protein 37A).

GN ZNF37A.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI\_TaxID=9606;

Qy 1

RP SEQUENCE FROM N.A.

RC TISSUE-Lung;

RA Strauberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RC TISSUE-Lung;

RA Strauberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RC TISSUE-Lung;

RA Strauberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RC TISSUE-Lung;

RA Strauberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RC TISSUE-Lung;

RA Strauberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-Berkeley;

RA Ashburner M., Mera S., Roote J., Lewis S.E., Blazet R., Davis T.,

RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,

RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,

RA Palazolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitehead K.,

RA Celinker S., Rubin G.M.;

RT "An exploration of the sequence of a 2.9-Mb region of the genome of

RT Drosophila melanogaster: the Adh region.";

RL Genetics 153:179-219 (1999).

RP SEQUENCE FROM N.A.

RC STRAIN-Berkeley;

RA Celinker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazet R.G.,

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman W.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Baller R.W.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,

RA Borokov D., Botchan M.R., Bouck P., Brockstein P., Brotter P.,

RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Fostler C., Gabrielian A.E., Garg N.S., Gehart W.M., Glasser K.,

RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,

RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson P.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Peled J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard D., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,

RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskaas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195 (2000).

RP SEQUENCE FROM N.A.

RC STRAIN-Berkeley;

RA Celinker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,

RA Evans C.A., Gocayne J.D., Amaralides P.G., Brandon R.C., Rogers Y.,

RA Banzon J., An H., Baldwin D., Banzon R.K.Y., Busam D.A.,

RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,

DR	Pfam; PF01352; KRA6; 1.
DR	Pfam; PF00096; zf_C2H2; 16.
DR	PRINTS; PR00048; ZINC_FINGER.
DR	ProDom; PD000003; ZnF_C2H2; 11.
DR	SMART; SM00349; KRA6; 1.
DR	SMART; SM00355; ZnF_C2H2; 16.
DR	PROSITE; PS00805; KRA6; 1.
DR	PROSITE; PS00282; ZINC_FINGER_C2H2_1; 16.
DR	PROSITE; PS0157; ZINC_FINGER_C2H2_2; 16.
KW	Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ	SEQUENCE 608 AA; 70548 MW; 4EF6B01D2730A4E CRC64;
OY	Query Match
DB	Best Local Similarity 75.5%; Score 105; DB 11; Length 608;
	Matches 19; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
	1 YQCEICGKSPSDKSNLTRHRIRITG 25
	354 YQCCECGKSPSQLCNLTFRHRIHTG 378
	RESULT 5
	Q8BKW8
ID	Q8BKW8 PRELIMINARY; PRT; 608 AA.
AC	Q8BKW8;
DT	01-MAR-2003 (TREMBLrel. 23, Created)
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE	Zinc finger protein 189 homolog.
OS	Mus musculus (Mouse).
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCHI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Body;
RX	MEDLINE=22354683; PubMed=12466851;
RA	The FANTOM Consortium,
RT	the RIKEN Genome Exploration Research Group Phase I & II Team;
RL	"Analysis of the mouse transcriptome based on functional annotation of
RL	60,770 full-length cDNAs." ;
DR	Nature 420:563-573(2002).
EMBL	AK049475; BAC33769.1; -
SQ	SEQUENCE 608 AA; 70522 MW; SC07ACT1C8PB3F06 CRC64;
OY	Query Match
DB	Best Local Similarity 75.5%; Score 105; DB 11; Length 608;
	Matches 19; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
	1 YQCEICGKSPSDKSNLTRHRIRITG 25
	354 YQCCECGKSPSQLCNLTFRHRIHTG 378
	RESULT 6
	Q8BKP2
ID	Q8BKP2 PRELIMINARY; PRT; 608 AA.
AC	Q8BKP2;
DT	01-MAR-2003 (TREMBLrel. 23, Created)
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE	Zinc finger protein 189 homolog.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCHI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX	MEDLINE=22354683; PubMed=12466851;
RA	The FANTOM Consortium,
RT	the RIKEN Genome Exploration Research Group Phase I & II Team;
TL	"Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 RC EMBL; AK051265; BAC34582.1; -  
 SQ SEQUENCE 608 AA; 70522 MW; 862F0214A16BA36F CRC64;

Query Match 75.5%; Score 105; DB 11; Length 608;  
 Best Local Similarity 76.0%; Pred. No. 9.8e-08;  
 Matches 19; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 YQCEICGKSPDKNLTHRLRIHTG 25  
 DB 354 YQCEICGKSPDKNLTHRLRIHTG 378

## RESULT 7

Q9R2Q3 PRELIMINARY; PRT; 566 AA.

AC Q9R2Q3; 01-JUN-2002 (TRENBLrel. 21, Created)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE Similar to KIAA1948 protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strauberg R.;  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC027344; AAH27344.1; -  
 DR InterPro; IPR001909; KRAB.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR Pfam; PF01352; KRAB.1;  
 DR Pfam; PF00096; Zf-C2H2; 10.  
 DR SMART; SM00355; KRAB.1;  
 DR SMART; SM00355; Znf\_C2H2; 10.  
 DR PROSITE; PS00805; KRAB.1;  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 10.  
 DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 10.  
 KW Metal-binding; Zinc; Zinc-finger.  
 SQ SEQUENCE 566 AA; 64569 MW; CC06DC1CE512056D CRC64;

Query Match 74.8%; Score 104; DB 11; Length 566;  
 Best Local Similarity 76.0%; Pred. No. 1.3e-07;  
 Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 YQCEICGKSPDKNLTHRLRIHTG 25  
 DB 251 YQCEICGKSPDKNLTHRLRIHTG 275

## RESULT 8

Q96N22 PRELIMINARY; PRT; 576 AA.

AC Q96N22; 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE Hypothetical protein FLJ31526.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nimomiyu K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,  
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,  
 RA Katsura N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
 RA Matsuo K., Nakamura Y., Sekine W., Kikuchi H., Murakawa K.,  
 RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,

RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.,  
 RT "NEO human cDNA sequencing project."  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; AK056088; BAB71090.1; -  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR InterPro; IPR007086; Znf\_C2H2\_sub.  
 DR Pfam; PF00096; Zf-C2H2; 15.  
 DR PRINTS; PR00048; ZINC\_FINGER.  
 DR PRODOM; PD000003; Znf\_C2H2; 2.  
 DR SMART; SM00355; Znf\_C2H2; 16.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 10.  
 DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 16.  
 KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;  
 KW Zinc-finger.  
 SQ SEQUENCE 576 AA; 66620 MW; 4396672D34B99CD CRC64;

Query Match 74.8%; Score 104; DB 4; Length 576;  
 Best Local Similarity 72.0%; Pred. No. 1.3e-07;  
 Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 YQCEICGKSPDKNLTHRLRIHTG 25  
 DB 493 YQCEICGKSPDKNLTHRLRIHTG 517

## RESULT 9

Q9ULI4 PRELIMINARY; PRT; 606 AA.

AC Q9ULI4; 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE D25J06.2 (Novel zinc finger protein) (Fragment).  
 GN D25J06.2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Williams S.;  
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; Z84476; CAB55432.2; -  
 DR HSRP; P08047; 1982.  
 DR InterPro; IPR001909; KRAB.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR InterPro; IPR007086; Znf\_C2H2\_sub.  
 DR Pfam; PF01352; KRAB.1;  
 DR Pfam; PF00096; Zf-C2H2; 14.  
 DR PRINTS; PR00048; ZINC\_FINGER.  
 DR SMART; SM00348; KRAB.1;  
 DR SMART; SM00355; Znf\_C2H2; 14.  
 DR PROSITE; PS00805; KRAB.1;  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 14.  
 DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 14.  
 KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.  
 FT NON TER 1  
 FT 606  
 SQ SEQUENCE 606 AA; 69975 MW; B2C611019C7E3B8 CRC64;

Query Match 74.8%; Score 104; DB 4; Length 606;  
 Best Local Similarity 68.0%; Pred. No. 1.4e-07;  
 Matches 17; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 YQCEICGKSPDKNLTHRLRIHTG 25  
 DB 334 YQCEICGKSPDKNLTHRLRIHTG 358

## RESULT 10

Q8NA42

```

ID 08N42; PRELIMINARY; PRT; 475 AA.
AC 08N42;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein FLJ35863.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_Taxid=9606;
OK NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Houta T., Hiraoka S., Murakawa K., Takiguchi S., Kuseno J., Chiba Y.,
RA Matnabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie K., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto T., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Magatsuna M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isegai T.;
RT "NEMO human cDNA sequencing project."
RL Submitted (JUN-2002) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AK093182; BAC04086.1; -.
DR InterPro: IPR001909; KRA6.
DR InterPro: IPR007087; Znf_C2H2.
DR Pfam; PF00352; KRA6.1; 10.
DR Pfam; PF00096; Zf-C2H2; 10.
DR PRINTS; PR00048; ZINC_FINGER.
DR SMART; SM00349; KRA6.1.
DR SMART; SM00355; Znf_C2H2; 11.
DR PROSITE; PS00805; KRA6.1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 11.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 11.
DR Hypothetical protein; Metal-binding; Nuclear protein; Zinc;
KW Zinc-finger.
SQ
SEQUENCE 475 AA; 54613 MW; 2F47070F5C2E582B CRC64;

Query Match 74.1%; Score 103; DB 4; Length 475;
Best Local Similarity 72.0%; Pred. No. 1.6e-07;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSFSDKSNLTRHLRIHTG 25
Db 422 YECNEGKAFKFCNSLTRLRIHTG 446

RESULT 11
Q9UFIH PRELIMINARY; PRT; 106 AA.
AC Q9UFIH;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_Taxid=9606;
OK NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Poustka A., Klein M., Neues H.W., Gaassenhuber J., Wiemann S.;
RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AL122078; CAB59253.1; -.
DR InterPro: IPR007087; Znf_C2H2.
DR InterPro: IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; Zf-C2H2; 3.
DR PRINTS; PR00048; ZINC_FINGER.

```

```

DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;
KW Zinc-finger.
FT NON_TER
SQ
SEQUENCE 106 AA; 12095 MW; 3D448E0C94F5F16B CRC64;

Query Match 73.4%; Score 102; DB 4; Length 106;
Best Local Similarity 64.0%; Pred. No. 4.8e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSFSDKSNLTRHLRIHTG 25
Db 26 YKCEVCGKSFSDKSNLTRHLRIHTG 50

RESULT 12
Q8R0KO PRELIMINARY; PRT; 337 AA.
AC Q8R0KO;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to unnamed protein product.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC026676; AAH26676.1; -.
DR InterPro: IPR007087; Znf_C2H2.
DR Pfam; PF00096; Zf-C2H2; 8.
DR SMART; SM00355; Znf_C2H2; 8.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 8.
KW Metal-binding; Zinc; Zinc-finger.
SQ
SEQUENCE 337 AA; 37455 MW; C0094DA11A98C312 CRC64;

Query Match 73.4%; Score 102; DB 11; Length 337;
Best Local Similarity 64.0%; Pred. No. 1.6e-07;
Matches 16; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 YQCEICGKSFSDKSNLTRHLRIHTG 25
Db 251 YRCDVCGKAFSDCSALVRLRVHSG 275

RESULT 13
Q8CSG3 PRELIMINARY; PRT; 496 AA.
AC Q8CSG3;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Finger protein ZNF50 homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Gonad;
RA MEDLINE-22354683; PubMed=12466851;
RL The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."

```

RL Nature 420:563-573(2002).  
 DR EMBL; AK078581; BAC7338.1; -  
 SQ SEQUENCE 496 AA; 55416 MW; 329A26F5AD496026 CRC64;

Query Match 73.4%; Score 102; DB 11; Length 496;  
 Best Local Similarity 64.0%; Pred. No. 2.3e-07;  
 Matches 16; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDCKSNLTFRHRIHTG 25  
 DB 410 YRCVCGKAFSDCSALVRLRHVHSG 434

## RESULT 14

Q8BGS5 PRELIMINARY; PRT; 496 AA.

AC Q8BGS5 01-MAR-2003 (TREMblrel. 23, Created)  
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Finger protein ZNF50 homolog.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCB1; TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Stomach;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The PANTOM Consortium,  
 RT "The Riken Genome Exploration Research Group Phase I & II Team;  
 RA "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK052559; BAC35038.1; -  
 DR EMBL; AK052507; BAC39388.1; -  
 SQ SEQUENCE 496 AA; 55458 MW; 74E7EB81A03D642E CRC64;

Query Match 73.4%; Score 102; DB 11; Length 496;  
 Best Local Similarity 64.0%; Pred. No. 2.3e-07;  
 Matches 16; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDCKSNLTFRHRIHTG 25  
 DB 410 YRCVCGKAFSDCSALVRLRHVHSG 434

## RESULT 15

Q8NA64 PRELIMINARY; PRT; 530 AA.

AC Q8NA64 01-OCT-2002 (TREMblrel. 22, Created)  
 DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Hypothetical protein FLJ35804 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NC NCB1; TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA Oshita A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,  
 RA Arita M., Muraishi K., Yuuki H., Hara H., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuna M.,  
 RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,  
 RA Sugano S., Nagahori K., Masuho Y., Nagai K., Isogai T.,  
 RT "NEDO human cDNA sequencing project";  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; AK093123; BAC04064.1; -

DR InterPro; IPR001909; KRAB.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR InterPro; IPR007086; Znf\_C2H2\_sub.  
 DR Pfam; PF01352; KRAB; 1.  
 DR Pfam; PR00096; ZF-C2H2; 13.  
 DR PRINTS; PR00048; ZINCFINGER.  
 DR SMART; SM00349; KRAB; 1.  
 DR SMART; SM00355; ZNF\_C2H2; 13.  
 DR PROSITE; PS50805; KRAB; 1.  
 DR PROSITE; PS50028; ZINC\_FINGER\_C2H2\_1; 13.  
 DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 13.  
 KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;  
 KW Zinc-finger.

FT NON\_TER 530 530  
 SQ SEQUENCE 530 AA; 61275 MW; 9491D3767BF5C5E4 CRC64;

Query Match 73.4%; Score 102; DB 4; Length 530;  
 Best Local Similarity 68.0%; Pred. No. 2.5e-07;  
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDCKSNLTFRHRIHTG 25  
 DB 438 YACTVCGKAFSOKSNLTFRHRIHTG 462

Search completed: January 6, 2004, 16:11:51  
 Job time : 37 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: January 6, 2004, 16:05:21 ; Search time 11 Seconds  
(without alignments)

106.879 Million cell updates/sec

Title: US-09-774-940A-1

Perfect score: 139  
Sequence: 1 YCCETGKSGFSFKSNLTRLRHTG 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	106	76.3	242 1 Z37A_HUMAN	P17032 homo sapien
2	104	74.8	626 1 Z189_HUMAN	O75820 homo sapien
3	102	73.4	803 1 Z226_HUMAN	Q9HYT6 homo sapien
4	102	73.4	913 1 Z228_HUMAN	Q9YJUS homo sapien
5	101	72.7	485 1 Z165_HUMAN	P49910 homo sapien
6	101	72.7	576 1 Z431_HUMAN	O8TF32 homo sapien
7	100	71.9	247 1 Z273_HUMAN	O14593 homo sapien
8	100	71.9	595 1 ZN85_HUMAN	Q05923 homo sapien
9	100	71.9	1191 1 ZN91_HUMAN	Q05481 homo sapien
10	99	71.2	348 1 Z134_HUMAN	P52741 homo sapien
11	98	70.5	169 1 ZF58_MOUSE	P16372 mus musculu
12	98	70.5	473 1 ZF38_HUMAN	O9Y5A6 homo sapien
13	98	70.5	474 1 Z256_HUMAN	O9Y2P7 homo sapien
14	98	70.5	555 1 ZF38_MOUSE	O07231 mus musculu
15	98	70.5	604 1 GLAS_DROME	P13360 drosophila
16	98	70.5	683 1 Z263_HUMAN	O14978 homo sapien
17	97	69.8	406 1 Z332_HUMAN	O961W9 homo sapien
18	97	69.8	810 1 Z33A_HUMAN	Q06730 homo sapien
19	96	69.1	216 1 Z434_HUMAN	O9XK65 homo sapien
20	96	69.1	275 1 Z253_HUMAN	O9Y346 homo sapien
21	96	69.1	325 1 ZN36_HUMAN	P17029 homo sapien
22	96	69.1	367 1 Z211_HUMAN	Q13398 homo sapien
23	96	69.1	424 1 ZFPI_MOUSE	P08042 mus musculu
24	96	69.1	469 1 Z135_HUMAN	P52742 homo sapien
25	96	69.1	489 1 ZN71_HUMAN	Q9NQZ8 homo sapien
26	96	69.1	574 1 YF73_HUMAN	O99255 homo sapien
27	96	69.1	643 1 ZN23_HUMAN	P17038 homo sapien
28	96	69.1	803 1 ZN43_HUMAN	P17038 homo sapien
29	95	68.3	379 1 Z11B_HUMAN	Q05732 homo sapien
30	95	68.3	428 1 ZN83_HUMAN	P51522 homo sapien
31	95	68.3	738 1 ZN25_HUMAN	P17030 homo sapien
32	94	67.6	56 1 ZNF_BOVIN	Q28151 bos taurus
33	94	67.6	292 1	

34	94	67.6	292 1	OZF_HUMAN	Q15072 homo sapien
35	94	67.6	301 1	Z11A_HUMAN	P17013 homo sapien
36	94	67.6	394 1	Z200_HUMAN	P98182 homo sapien
37	94	67.6	457 1	Z140_HUMAN	P52738 homo sapien
38	94	67.6	534 1	Z397_HUMAN	O8NF99 homo sapien
39	94	67.6	616 1	ZN93_HUMAN	P35789 homo sapien
40	94	67.6	671 1	Z282_HUMAN	Q9UDV7 homo sapien
41	94	67.6	734 1	ZN42_HUMAN	P28698 homo sapien
42	94	67.6	783 1	ZF25_HUMAN	Q9U115 homo sapien
43	94	67.6	839 1	Z347_HUMAN	Q96SE7 homo sapien
44	93	66.9	347 1	ZFPI_MOUSE	P08043 mus musculu
45	93	66.9	364 1	Z430_HUMAN	Q9H9L1 homo sapien

## ALIGNMENTS

RESULT 1  
ID Z37A\_HUMAN STANDARD; PRT; 242 AA.  
AC P17032;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Zinc finger protein 37A (Zinc finger protein KOX21) (Fragment).  
GN ZNF37A OR ZNF37 OR KOX21.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9329119; PubMed=8464732;  
RA Tunnaciliffe A., Liu L., Moore J.K., Leverha M.A., Jackson M.S.,  
RA Ferguson-Smith M.A., Thiesen H.-J., Ponder B.A.;  
RT "Duplicated KOX zinc finger gene clusters flank the centromere of  
RT human chromosome 10: evidence for a pericentric inversion during  
RT primate evolution";  
RL Nucleic Acids Res. 21:1409-1417(1993).  
RN [2]  
RP SEQUENCE OF 8-63 FROM N.A.  
RC TISSUE=Lymphoid;  
RX MEDLINE=91145339; PubMed=2288909;  
RA Thiesen H.-J.;  
RT "Multiple genes encoding zinc finger domains are expressed in human T  
RT cells";  
RL New Biol. 2:363-374(1990).  
CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
CC FINGER PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL, X69115; CA44886.1; -;  
DR EMBL, X52352; CA43578.1; -;  
DR PIR, S30238; S30238.  
DR HSSP, P08046; 1AIG.  
DR Genew; HGNC:13102; ZNF37A.  
DR GO, GO:0005634; C:nucleus; NAS.  
DR GO, GO:0003700; F:transcription factor activity; NAS.  
DR GO, GO:0006355; P:regulation of transcription, DNA-dependent; NAS.  
DR InterPro, IPR007087; Znf\_C2H2.  
DR Pfam, PF00096; Znf\_C2H2; 8.  
DR SMART, SM00355; ZNF\_C2H2; 8.  
DR PROSITE, PS00028; ZINC\_FINGER\_C2H2\_1; 8.  
DR PROSITE, PS50157; ZINC\_FINGER\_C2H2\_2; 8.

KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;  
KW Nuclear protein; Repeat.  
FT ZN\_FING 1 30 C2H2-TYPE.  
FT ZN\_FING 8 58 C2H2-TYPE.  
FT ZN\_FING 36 58 C2H2-TYPE.  
FT ZN\_FING 64 86 C2H2-TYPE.  
FT ZN\_FING 92 114 C2H2-TYPE.  
FT ZN\_FING 120 142 C2H2-TYPE.  
FT ZN\_FING 148 170 C2H2-TYPE.  
FT ZN\_FING 176 198 C2H2-TYPE.  
FT ZN\_FING 204 226 C2H2-TYPE.  
SQ SEQUENCE 242 AA; 27810 MM; 281045182ASDAD8 CRC64;  
Query Match 76.3%; Score 106; DB 1; Length 242;  
Best Local Similarity 72.0%; Pred. No. 1.4e-08;  
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
OY 1 YOCCKGKSPSKSNLTRLRHRIHTG 25  
Db 92 YECNCKGKSPSKSNLTRLRHRIHTG 116  
RESULT 2  
ID 2189 HUMAN STANDARD; PRT; 626 AA.  
AC 075820: 075820; Q9UBI4; Q9UPF9; Q9UPF0; Q9UPF1;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Zinc finger protein 189.  
OS ZNF189.  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4).  
RC TISSUE=Bone marrow;  
RX MEDLINE=98317535; Pubmed=9653648;  
RA Odeberg J., Roessck O., Gudmundsson G., Ahmadian A., Roshani L.,  
RA Williams C., Larsson C., Ponten F., Uhlen M., Asheim H.-C.,  
RA Lundberg J.,  
RT "Cloning and characterization of ZNF189, a novel human Kruppel-like  
RT zinc finger gene localized to chromosome 9q22-q31."  
RL Genomics 50:213-221 (1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 4).  
RX MEDLINE=99346145; Pubmed=10415338;  
RA Odeberg J., Ahmadian A., Williams C., Uhlen M., Ponten F.,  
RA Lundberg J.,  
RT "Context-dependent Tag-polymerase-mediated nucleotide alterations, as  
RT revealed by direct sequencing of the ZNF189 gene: implications for  
RT mutation detection."  
RL Gene 235:103-109 (1999).  
CC -1- FUNCTION: May function as a transcription factor.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=1;  
CC IsoId=O75820-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=O75820-2; Sequence=VSP\_006900;  
CC Name=3; Synonyms=B2;  
CC IsoId=O75820-3; Sequence=VSP\_006901;  
CC Name=4;  
CC IsoId=O75820-4; Sequence=VSP\_006899;  
CC -1- TISSUE SPECIFICITY: Ubiquitous.  
CC -1- SIMILARITY: BELONGS TO THE KRUPPEL FAMILY OF C2H2-TYPE ZINC-  
CC FINGER PROTEINS.  
CC -1- SIMILARITY: Contains 1 KRAE domain.  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U95992; AAC25910.1; -;  
DR EMBL; U95991; AAC25909.1; -;  
DR EMBL; U75454; AAC39798.1; -;  
DR EMBL; AF025770; AAC39799.1; -;  
DR EMBL; AF025771; AAC39800.1; -;  
DR EMBL; AF025772; AAC50527.1; -;  
DR HSSP; P08047; 1SP2.  
DR GeneW; HGNC:12980; ZNF189.  
DR MIM; 603132; -;  
DR GO; GO:0003700; F:transcription factor activity; TAS.  
DR GO; GO:000122; F:negative regulation of transcription from P...; TAS.  
DR InterPro; IPR001909; KRAE.  
DR InterPro; IPR007087; Znf\_C2H2.  
DR InterPro; IPR007086; Znf\_C2H2\_sub.  
DR Pfam; PF01352; KRAE; 1.  
DR Pfam; PF00096; ZF-C2H2; 16.  
DR PRINTS; PR00048; ZINC\_FINGER.  
DR ProDom; PD000003; Znf\_C2H2; 11.  
DR SMART; SM00349; KRAE; 1.  
DR SMART; SM00355; Znf\_C2H2; 16.  
DR PROSITE; PS00805; KRAE; 1.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 16.  
DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 16.  
KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;  
KW Nuclear protein; Repeat; Alternative splicing.  
FT ZN\_FING 14 89 KRAE.  
FT ZN\_FING 148 170 C2H2-TYPE.  
FT ZN\_FING 176 198 C2H2-TYPE.  
FT ZN\_FING 204 226 C2H2-TYPE.  
FT ZN\_FING 232 254 C2H2-TYPE.  
FT ZN\_FING 260 282 C2H2-TYPE.  
FT ZN\_FING 288 310 C2H2-TYPE.  
FT ZN\_FING 316 338 C2H2-TYPE.  
FT ZN\_FING 344 366 C2H2-TYPE.  
FT ZN\_FING 372 394 C2H2-TYPE.  
FT ZN\_FING 400 422 C2H2-TYPE.  
FT ZN\_FING 456 478 C2H2-TYPE.  
FT ZN\_FING 484 506 C2H2-TYPE.  
FT ZN\_FING 512 534 C2H2-TYPE.  
FT ZN\_FING 540 562 C2H2-TYPE.  
FT ZN\_FING 568 590 C2H2-TYPE.  
FT ZN\_FING 599 621 C2H2-TYPE.  
FT VASAPLIC 1 42 Missing (in isoform 4).  
FT VASAPLIC 12 25 Missing (in isoform 2).  
FT VASAPLIC 1 95 Missing (in isoform 3).  
FT VASAPLIC 1 95 Missing (in isoform 3).  
FT CONFLICT 588 588 Missing (in isoform 2).  
FT CONFLICT 588 588 Missing (in isoform 3).  
FT CONFLICT 588 588 Missing (in isoform 3).  
SQ SEQUENCE 626 AA; 72978 MM; C1100F9943F15B6 CRC64;  
Query Match 74.8%; Score 104; DB 1; Length 626;  
Best Local Similarity 76.0%; Pred. No. 7.3e-08;  
Matches 19; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
OY 1 YOCCKGKSPSKSNLTRLRHRIHTG 25  
Db 372 YOCCKGKSPSKSNLTRLRHRIHTG 396  
RESULT 3  
ID 2226 HUMAN STANDARD; PRT; 803 AA.  
AC Q9NYT6; Q96TB6; Q9NS44;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)



DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Zinc finger protein 226.  
 GN ZNF226.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shannon M., Branscomb E., Hauser L., Gordon L., Ashworth L.,  
 RT Stubbs L.;  
 RT "Differential expansion of homologous zinc-finger gene families in  
 RT human chromosome 19q13.2 and mouse chromosome 7.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Kodyiam V., Ge Y., Severin J., Krummel G.K., Grable L.,  
 RA Kvikstad E., Gordon L., Shannon M., Brower A., Olsen A.S., Smith L.M.;  
 RT "Sequence analysis of a 1Mb region in 19q13.2 containing a zinc finger  
 RT gene cluster.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Watanabe M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakami K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishi S., Kawai Y., Saito K.,  
 RA Yamamoto J., Makamatsu A., Nakamura Y., Nagahara K., Masuko Y.,  
 RA Nimochi K., Iwayanagi T.;  
 RT "WED human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Tissue=Muscle;  
 RC MEDLINE=22388257; PubMed=12477932;  
 RX Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner F.S., Shemen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marutka K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stalcen M., Soares M.B., Bonaldo M.F., Caravita T.L., Scheetz T.E.,  
 RA Brownstein M.T., Usdin E.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield V.S.N., Krzywinski M.I., Skalski U., Small D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -1- SIMILARITY: Contains 1 KRAB domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL: AF28419; AAF3766.1;  
 DR EMBL: AC074331; AAF8103.1;  
 DR EMBL: AK023091; BAB14398.1;

DR EMBL: BC024197; AAH24197.1;  
 DR HSSP: P08047; ISP2.  
 DR Genew: HGNC:13019; ZNF226.  
 DR InterPro: IPR001909; KRAB.  
 DR InterPro: IPR007087; ZNF\_C2H2.  
 DR Pfam: PF01352; KRAB.1.  
 DR Pfam: PF00066; ZF-C2H2.18.  
 DR Pfam: PD00003; ZNF\_C2H2.15.  
 DR SMART: SM00349; KRAB.1.  
 DR SMART: SM00355; ZNF\_C2H2.18.  
 DR PROSITE: PS00805; KRAB.1.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 17.  
 DR PROSITE: PS0157; ZINC\_FINGER\_C2H2\_2; 19.  
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;  
 KW Nuclear protein.  
 FT DOMAIN 8  
 FT ZN\_FING 252 274 KRAB.  
 FT ZN\_FING 307 329 C2H2-TYPE (DEGENERATE).  
 FT ZN\_FING 335 357 C2H2-TYPE.  
 FT ZN\_FING 363 385 C2H2-TYPE.  
 FT ZN\_FING 391 413 C2H2-TYPE.  
 FT ZN\_FING 419 441 C2H2-TYPE.  
 FT ZN\_FING 447 469 C2H2-TYPE.  
 FT ZN\_FING 475 497 C2H2-TYPE.  
 FT ZN\_FING 503 525 C2H2-TYPE.  
 FT ZN\_FING 531 553 C2H2-TYPE.  
 FT ZN\_FING 559 581 C2H2-TYPE.  
 FT ZN\_FING 587 609 C2H2-TYPE.  
 FT ZN\_FING 615 637 C2H2-TYPE.  
 FT ZN\_FING 643 665 C2H2-TYPE.  
 FT ZN\_FING 671 693 C2H2-TYPE.  
 FT ZN\_FING 699 721 C2H2-TYPE.  
 FT ZN\_FING 727 749 C2H2-TYPE.  
 FT ZN\_FING 755 777 C2H2-TYPE.  
 FT CONFLICT 195 195 K -> E (IN REF. 1).  
 SQ SEQUENCE 803 AA; 91920 MW; 256DE7551A932A8 CRC64;  
 Query Match 73.4%; Score 102; DB 1; Length 803;  
 Best Local Similarity 76.0%; Pred. No. 1.9e-07;  
 Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 YQCEICGKSPDSDXNTRHRIHTG 25  
 Db 755 YKCEICGKSPDSDXNTRHRIHTG 779  
 ID 2228\_HUMAN STANDARD; PRT; 913 AA.  
 AC Q9UTJ3; Q9HCA7;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Zinc finger protein 228.  
 GN ZNF228.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shannon M., Branscomb E., Hauser L., Gordon L., Ashworth L.,  
 RT Stubbs L.;  
 RT "Differential expansion of homologous zinc-finger gene families in  
 RT human chromosome 19q13.2 and mouse chromosome 7.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Kodyiam V., Ge Y., Krummel G.K., Grable L.,  
 RA Kvikstad E., Gordon L., Shannon M., Brower A., Olsen A.S., Smith L.M.;  
 RT "Sequence analysis of a 1Mb region in 19q13.2 containing a zinc finger  
 RT gene cluster.";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

```

CC - FUNCTION: May function as a transcription factor.
CC - SUBCELLULAR LOCATION: Nuclear (Probable).
CC - SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC - SIMILARITY: Contains 1 KRAB domain.
CC - SIMILARITY: Contains 17 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF198358; AAF12816.1; -
DR EMBL; AC084239; AAG33968.1; -
DR HSSP; P08047; ISP2; C2H2_1; -
DR Genew; HGNC:13021; ZNF228.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 13.
DR PRINTS; PR00048; ZINC_FINGER.
DR PRODOM; PD000003; Znf_C2H2; 12.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 13.
DR PROSITE; PS50805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 17.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein.
FT DOMAIN 8 79
FT ZN_FING 258 280 KRAB, C2H2-TYPE 1 (DEGENERATE).
FT ZN_FING 443 465 C2H2-TYPE 2 (DEGENERATE).
FT ZN_FING 471 493 C2H2-TYPE 3 (DEGENERATE).
FT ZN_FING 497 519 C2H2-TYPE 4 (DEGENERATE).
FT ZN_FING 525 547 C2H2-TYPE 5.
FT ZN_FING 553 575 C2H2-TYPE 6.
FT ZN_FING 581 603 C2H2-TYPE 7.
FT ZN_FING 609 631 C2H2-TYPE 8.
FT ZN_FING 637 659 C2H2-TYPE 9.
FT ZN_FING 665 687 C2H2-TYPE 10.
FT ZN_FING 693 715 C2H2-TYPE 11.
FT ZN_FING 721 743 C2H2-TYPE 12.
FT ZN_FING 749 771 C2H2-TYPE 13.
FT ZN_FING 777 799 C2H2-TYPE 14.
FT ZN_FING 805 827 C2H2-TYPE 15.
FT ZN_FING 833 855 C2H2-TYPE 16.
FT ZN_FING 861 883 C2H2-TYPE 17.
FT CONFILCT 3 5 VSK -> KFG (IN REF. 2).
FT CONFILCT 113 113 T -> A (IN REF. 2).
FT CONFILCT 143 143 E -> V (IN REF. 2).
FT CONFILCT 222 222 E -> K (IN REF. 2).
FT CONFILCT 261 261 S -> T (IN REF. 2).
FT CONFILCT 392 392 S -> N (IN REF. 2).
FT CONFILCT 446 446 Q -> E (IN REF. 2).
FT CONFILCT 485 485 H -> Y (IN REF. 2).
FT CONFILCT 813 813 A -> G (IN REF. 2).
SQ SEQUENCE 913 AA; 105702 MW; 8D19F5D27CD9FA6B CRC64;

```

Query Match 73.4%; Score 102; DB 1; Length 913;  
 Best Local Similarity 64.0%; Pred. No. 2.2e-07;  
 Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDSKSNLTRHLRIHTG 25  
 Db 833 YKCEVCGKGFQSRNSNLAHORVHTG 857

RESULT 5  
 Z165\_HUMAN

```

ID Z165_HUMAN STANDARD; PRT; 485 AA.
AC P49910;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 165 (ZNF165).
GN ZNF165 OR ZP165.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBT_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96039260; PubMed=7490084;
RA Tinsworth K.N., Dyane A., Jones M., Affara N.A.;
RT "Characterization of a novel zinc finger gene (ZNF165) mapping to
RT 6p21 that is expressed specifically in testis.";
RL Genomics 28:485-490(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=97386587; PubMed=9244436;
RA Lee P.L., Gelbart T., West C., Adams M., Blackstone R., Beutler E.;
RT "Three genes encoding zinc finger proteins on human chromosome 6p21.3:
RT members of a new subclass of the Kruppel gene family containing the
RT conserved SCAN box domain.";
RL Genomics 43:191-201(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC - SUBCELLULAR LOCATION: Nuclear (Potential).
CC - TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN TESTIS.
CC - SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC - SIMILARITY: Contains 1 SCAN box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X84801; CAAS9268.1; -
DR EMBL; U78722; AAC51658.1; -
DR EMBL; U88086; AAD04755.1; -
DR EMBL; U88084; AAD04755.1; JOINED.
DR EMBL; U88085; AAD04755.1; JOINED.
DR EMBL; AL121944; CAB82080.1; -
DR PIR; S52411; S52411.
DR HSSP; P07248; IARD.
DR Genew; HGNC:12953; ZNF165.
DR MIM; 600834; -
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; NAS.
DR InterPro; IPR003309; Treg_SCAN.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF02023; SCAN; 1.
DR Pfam; PF00096; zf-C2H2; 6.
DR PRODOM; PD000003; Znf_C2H2; 4.
DR SMART; SM00431; IER; 1.
DR SMART; SM00355; Znf_C2H2; 5.
DR PROSITE; PS50804; SCAN_BOX; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 6.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT DOMAIN 62 127 SCAN BOX.

```

```

FT  ZN_PING 344 366 C2H2-TYPE.
FT  ZN_PING 372 394 C2H2-TYPE.
FT  ZN_PING 400 422 C2H2-TYPE.
FT  ZN_PING 428 450 C2H2-TYPE.
FT  ZN_PING 456 478 C2H2-TYPE.
SQ  SEQUENCE 485 AA; 55770 MW; 1265D859D1713A8 CRC64;

Query Match 72.7%; Score 101; DB 1; Length 485;
Best Local Similarity 68.0%; Pred. No. 1.6e-07;
Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICKGKSPFSDKSNLTRHLRIHTG 25
Db 372 YECNECGKSPAESDILTRHRIHTG 396

RESULT 6
ID 2431 HUMAN STANDARD; PRT; 576 AA.
AC O8TRF32; O8IWC4;
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 431.
GN ZNF431 OR KIAA1969.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN 1;
RX MEDLINE=21842142; PubMed=11653319;
RA Nagase T., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XXII.
RT The complete sequences of 50 new cDNA clones which code for large
RT proteins.";
RL DNA Res. 8:319-327 (2001).
RN 2;
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang C.F., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Usdin T.B., Toshiyuki S., Carninci P., Franke C.,
RA Raha S.S., McManus N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McManus P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shcherbko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16889-16903 (2002).
CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC PINER PROTEINS.
CC -1- SIMILARITY: Contains 1 KRAB domain.
CC -1- SIMILARITY: Contains 13 C2H2-type zinc fingers.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AB075849; BAB8555.1; ALT_INT.
DR EMBL; BC040506; AAAH0506.1; -.
DR GenBank; HGNC:20809; ZNF431.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 12.
DR PRINTS; PR00048; ZINCFINGER.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 12.
DR PROSITE; PS50805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 12.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 13.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT DOMAIN 35 106
FT ZN_FING 176 198 KRAB.
FT ZN_FING 204 226 C2H2-TYPE 1 (DEGENERATE).
FT ZN_FING 232 254 C2H2-TYPE 2.
FT ZN_FING 260 282 C2H2-TYPE 3.
FT ZN_FING 288 310 C2H2-TYPE 4.
FT ZN_FING 316 338 C2H2-TYPE 5.
FT ZN_FING 344 366 C2H2-TYPE 6.
FT ZN_FING 372 394 C2H2-TYPE 7.
FT ZN_FING 400 422 C2H2-TYPE 8.
FT ZN_FING 428 450 C2H2-TYPE 9.
FT ZN_FING 456 478 C2H2-TYPE 10.
FT ZN_FING 484 506 C2H2-TYPE 11.
FT ZN_FING 512 534 C2H2-TYPE 12.
FT ZN_FING 198 534 C2H2-TYPE 13.
FT CONFLICT 198 534 H -> R (IN REF. 2).
SQ SEQUENCE 576 AA; 67216 MW; 532774B69EC9E2A CRC64;

Query Match 72.7%; Score 101; DB 1; Length 576;
Best Local Similarity 64.0%; Pred. No. 1.9e-07;
Matches 16; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICKGKSPFSDKSNLTRHLRIHTG 25
Db 400 YKEVCGKAPNENSSNLTTHMCIHTG 424

RESULT 7
ID 2273 HUMAN STANDARD; PRT; 247 AA.
AC Q14593;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 273 (Zinc finger protein HZF9) (Fragment).
GN ZNF273.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RC MEDLINE=95169271; PubMed=7865130;
RA Ahrink M., Aveskog M., Hellman L.;
RT "Isolation of cDNA clones for 42 different Kruppel-related zinc finger
RT proteins expressed in the human monoblast cell line U-937.";
RL DNA Cell Biol. 14:125-136 (1995).
CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC PINER PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: X78932, CAA5532.1; -  
 DR PIR: S47070, S47070.  
 DR HSP: P08047, I5P2.  
 DR Genew: HGNC:13067; ZNF273.  
 DR MIM: 604756; -  
 DR InterPro: IPR007087; Znf\_C2H2, 6.  
 DR Pfam: PF00096; Zf\_C2H2, 7.  
 DR ProDom: PD000003; Znf\_C2H2, 3.  
 DR SMART: SM00355; Znf\_C2H2, 6.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1, 6.  
 DR PROSITE: PS50157; ZINC\_FINGER\_C2H2\_2, 9.  
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;  
 KW Nuclear protein; Repeat.  
 FT NON TER 1  
 FT ZN\_FING <1 19 C2H2-TYPE.  
 FT ZN\_FING 25 47 C2H2-TYPE.  
 FT ZN\_FING 53 75 C2H2-TYPE.  
 FT ZN\_FING 81 102 C2H2-TYPE (ATYPICAL).  
 FT ZN\_FING 108 130 C2H2-TYPE.  
 FT ZN\_FING 136 158 C2H2-TYPE.  
 FT ZN\_FING 164 186 C2H2-TYPE.  
 FT ZN\_FING 192 214 C2H2-TYPE.  
 FT ZN\_FING 220 242 C2H2-TYPE (ATYPICAL).  
 SQ SEQUENCE 247 AA; 28448 MW; 5AB556F46C334B2 CRC64;

Query Match 71.9%; Score 100; DB 1; Length 247;  
 Best Local Similarity 68.0%; Pred. No. 1.1e-07;  
 Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 YQCEICGKSFSDKSNLTRHRIHTG 25  
 Db 192 YKCECGKAFNRSSNLTFRKKIHTG 216

RESULT 8  
 ZN85\_HUMAN STANDARD; PRT; 595 AA.  
 AC Q03923;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Zinc finger protein 85 (Zinc finger protein HRF4) (HTF1).  
 GN ZNF85.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.  
 RC TISSUE=Placenta;  
 RX MEDLINE=98053537; PubMed=9839802;  
 RA Poncelet D.A., Bellefroid E.U., Bastiaens P.V., Demotie M.A.,  
 RA Marine J.C., Pendevel H., Alam Y., Devos N., Lecocq P.U., Ogawa T.,  
 RA Muller M., Martial J.A.;  
 RT "The evolutionarily conserved Kruppel-associated box domain defines a  
 RT subfamily of eukaryotic multifingered proteins";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:3608-3612(1991).  
 CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- TISSUE SPECIFICITY: EXPRESSED PREFERENTIALLY IN TESTICULAR  
 CC TISSUES.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY DURING EMBRYONIC DEVELOPMENT.  
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -1- SIMILARITY: Contains 1 KRAB domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: U35376; AAA79179.1; -  
 DR EMBL: M61668; AAA52689.1; -  
 DR EMBL: M61668; AAA58671.1; -  
 DR PIR: G02075; G02075.  
 DR HSP: P08048; ZNF.  
 DR TRANSFAC: T04990; -  
 DR Genew: HGNC:13160; ZNF85.  
 DR MIM: 603899; -  
 DR GO: GO:0005634; C:nucleus; TAS.  
 DR GO: GO:0003714; F:transcription co-repressor activity; TAS.  
 DR GO: GO:0003700; F:transcription factor activity; TAS.  
 DR InterPro: IPR001309; KRAB.  
 DR InterPro: IPR007087; Znf\_C2H2.  
 DR InterPro: IPR007086; Znf\_C2H2\_sub.  
 DR Pfam: PF00096; Zf\_C2H2, 15.  
 DR Pfam: PF01352; KRAB, 1.  
 DR PRINTS: PR00048; ZINC\_FINGER.  
 DR ProDom: PD000003; Znf\_C2H2, 13.  
 DR SMART: SM00349; KRAB, 1.  
 DR SMART: SM00355; Znf\_C2H2, 15.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1, 14.  
 DR PROSITE: PS50157; ZINC\_FINGER\_C2H2\_2, 15.  
 KW Transcription regulation; Zinc-finger; DNA-binding; Metal-binding;  
 KW Nuclear protein; Repeat; Repressor.  
 FT DOMAIN 4 75 KRAB.  
 FT ZN\_FING 146 168 C2H2-TYPE (DEGENERATE).  
 FT ZN\_FING 174 196 C2H2-TYPE.  
 FT ZN\_FING 202 224 C2H2-TYPE.  
 FT ZN\_FING 230 252 C2H2-TYPE.  
 FT ZN\_FING 258 280 C2H2-TYPE.  
 FT ZN\_FING 286 308 C2H2-TYPE.  
 FT ZN\_FING 314 336 C2H2-TYPE.  
 FT ZN\_FING 342 364 C2H2-TYPE.  
 FT ZN\_FING 370 392 C2H2-TYPE.  
 FT ZN\_FING 398 420 C2H2-TYPE.  
 FT ZN\_FING 426 448 C2H2-TYPE (DEGENERATE).  
 FT ZN\_FING 454 476 C2H2-TYPE.  
 FT ZN\_FING 482 504 C2H2-TYPE.  
 FT ZN\_FING 510 532 C2H2-TYPE.  
 FT ZN\_FING 538 560 C2H2-TYPE.  
 FT ZN\_FING 566 588 C2H2-TYPE.  
 FT ZN\_FING 584 606 R -> Q (IN REF. 2; AAA52689).  
 FT CONFLICT 84 84 R -> I (IN REF. 2; AAA58671).  
 FT CONFLICT 115 115 T -> R (IN REF. 2; AAA58671).  
 FT CONFLICT 177 177 G -> R (IN REF. 2; AAA58671).  
 FT CONFLICT 184 184 G -> R (IN REF. 2; AAA58671).  
 SQ SEQUENCE 595 AA; 68718 MW; 44AA0A236D62D43B CRC64;

Query Match 71.9%; Score 100; DB 1; Length 595;  
 Best Local Similarity 68.0%; Pred. No. 2.7e-07;  
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 YQCEICGKSFSDKSNLTRHRIHTG 25  
 Db 538 YKCECGKAFNRSSNLTFRKKIHTG 562

RESULT 9

ID	NAME	STANDARD	PRT	1191 AA.
DN91	HUMAN			
AC	005481			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Zinc finger protein 91 (Zinc finger protein HTRF1) (HPF1).			
GN	ZNF91.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RA	SEQUENCE FROM N.A.			
RP	MEDLINE=93223677; PubMed=8467795;			
RX	Bellefroid E.J., Poncelet D.A., Lecocq P.J., Revelant O.,			
RA	Martial U.A.;			
RT	"The evolutionarily conserved Kruppel-associated box domain defines a			
RL	subfamily of eukaryotic multifingered proteins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:3608-3612(1991).			
CC	-1- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	-1- SIMILARITY: BELONGS TO THE KRÜPPEL FAMILY OF C2H2-TYPE ZINC-			
CC	FINGER PROTEINS.			
CC	-1- SIMILARITY: Contains 1 KRAB domain.			
CC	-1- CAUTION: THE SEQUENCE FROM POSITION 1159 TO THE C-TERMINAL IS			
CC	DERIVED FROM THE TRANSLATION OF AN ALU REPEAT.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation.			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	-----			
EMBL	L11672; AA59469.1; ALT_SEQ.			
EMBL	M61871; AA58672.1; ALT_SEQ.			
DR	PIR; S35305; S35305.			
DR	HSSP; P08047; ISP2.			
DR	Genew; HGNC:13166; ZNF91.			
DR	MEM; 603971; -			
DR	GO; GO:0005634; C:nucleus; NAS.			
DR	GO; GO:0003700; E:transcription factor activity; NAS.			
DR	GO; GO:0008270; F:zinc ion binding activity; NAS.			
DR	GO; GO:0006355; P:regulation of transcription; NAS.			
DR	Interpro; IPR001909; KRAB.			
DR	Interpro; IPR007087; znf_C2H2.			
DR	Interpro; IPR007086; znf_C2H2_sub.			
DR	Pfam; PF01352; KRAB.1.			
DR	Pfam; PF00096; zf_C2H2; 34.			
DR	PRINTS; PR00048; ZINC_FINGER.			
DR	ProDom; PD000003; znf_C2H2; 20.			
DR	SMART; SM00349; KRAB.1.			
DR	SMART; SM00355; znf_C2H2; 33.			
DR	PROSITE; PS00805; KRAB.1.			
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 31.			
DR	PROSITE; PS01517; ZINC_FINGER_C2H2_2; 35.			
KW	Nuclear protein, Repeat.			
KW	Nuclear protein, Repeat.			
FT	DOMAIN 13 84	KRAB.		
FT	ZN_FING 154 176	C2H2-TYPE (DEGENERATE).		
FT	ZN_FING 179 200	C2H2-TYPE.		
FT	ZN_FING 208 232	C2H2-TYPE.		
FT	ZN_FING 238 260	C2H2-TYPE.		

```
CC CC      FT   ZN_FING     266    288    C2H2-TYPE.
CC CC      FT   ZN_FING     294    316    C2H2-TYPE.
CC CC      FT   ZN_FING     322    344    C2H2-TYPE.
CC CC      FT   ZN_FING     350    372    C2H2-TYPE.
CC CC      FT   ZN_FING     378    400    C2H2-TYPE.
CC CC      FT   ZN_FING     406    428    C2H2-TYPE.
CC CC      FT   ZN_FING     434    456    C2H2-TYPE.
CC CC      FT   ZN_FING     462    484    C2H2-TYPE.
CC CC      FT   ZN_FING     490    512    C2H2-TYPE.
CC CC      FT   ZN_FING     518    540    C2H2-TYPE.
CC CC      FT   ZN_FING     546    568    C2H2-TYPE.
CC CC      FT   ZN_FING     574    596    C2H2-TYPE.
CC CC      FT   ZN_FING     602    624    C2H2-TYPE.
CC CC      FT   ZN_FING     630    652    C2H2-TYPE.
CC CC      FT   ZN_FING     658    680    C2H2-TYPE.
CC CC      FT   ZN_FING     686    708    C2H2-TYPE.
CC CC      FT   ZN_FING     714    736    C2H2-TYPE.
CC CC      FT   ZN_FING     742    764    C2H2-TYPE.
CC CC      FT   ZN_FING     770    792    C2H2-TYPE.
CC CC      FT   ZN_FING     798    820    C2H2-TYPE.
CC CC      FT   ZN_FING     826    848    C2H2-TYPE.
CC CC      FT   ZN_FING     854    876    C2H2-TYPE.
CC CC      FT   ZN_FING     885    904    C2H2-TYPE.
CC CC      FT   ZN_FING     910    932    C2H2-TYPE.
CC CC      FT   ZN_FING     938    960    C2H2-TYPE.
CC CC      FT   ZN_FING     966    988    C2H2-TYPE.
CC CC      FT   ZN_FING     994    1016    C2H2-TYPE.
CC CC      FT   ZN_FING     1022    1044    C2H2-TYPE.
CC CC      FT   ZN_FING     1050    1072    C2H2-TYPE.
CC CC      FT   ZN_FING     1078    1100    C2H2-TYPE.
CC CC      FT   ZN_FING     1106    1128    C2H2-TYPE.
CC CC      FT   ZN_FING     1134    1156    C2H2-TYPE.
CC CC      SQ SEQUENCE     1191 AA; 137225 MW; 581056B1B8716D CRC64;

Query Match          71.9%; Score 100; DB 1; Length 1191;
Best Local Similarity 65.0%; Pred. No. 5,7e-07;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Cy       1 YQCEICGKSPFSDKSNTLRHLRIHTG 25
         ||| | | | | | | | | | | : | | |
Db       994 YKCCECGKAPSQSSTLTRHRMHMG 1018

RESULT 10
Z134_HUMAN STANDARD; PRT; 348 AA.
AC PS2741:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 134.
GN ZNF134.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
XX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Insulinoma;
RA MEDLINE=9604430; PubMed=7557990;
RX Tommerup N.; Vissing H.;
RT "Isolation and fine mapping of 16 novel human zinc finger-encoding
RT cDNAs identifiy putative candidate genes for developmental and
RT malignant disorders.";
RL Genomics 27:259-264(1995).
CC -1 FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -1 SUBCELLULAR LOCATION: Nuclear (potential).
CC -1 SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.
```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U09412; AACG0253.1; -  
 CC PIR: 138599; 138599.  
 CC HSSP: P08047; 1382.  
 CC Genew: HGNC:12918; ZNF134.  
 CC MIM: 604076; -  
 CC GO: GO:0005634; C:nucleus; NAS.  
 CC GO: GO:0003700; F:transcription factor activity; NAS.  
 CC GO: GO:000355; P:regulation of transcription, DNA-dependent; NAS.  
 CC InterPro: IPR007087; Znf\_C2H2.  
 CC Pfam: PF00096; zf-C2H2; 9.  
 CC ProDom: PD000003; Znf\_C2H2; 9.  
 CC SMART: SM00355; Znf\_C2H2; 9.  
 CC PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 9.  
 CC PROSITE: PS50157; ZINC\_FINGER\_C2H2\_2; 9.  
 CC Transcription regulation; Zinc-finger; DNA-binding; Metal-binding;  
 KW Nuclear protein; Repeat.  
 FT ZN\_FING 97 119 C2H2-TYPE.  
 FT ZN\_FING 125 147 C2H2-TYPE.  
 FT ZN\_FING 153 175 C2H2-TYPE.  
 FT ZN\_FING 181 203 C2H2-TYPE.  
 FT ZN\_FING 209 231 C2H2-TYPE.  
 FT ZN\_FING 237 259 C2H2-TYPE.  
 FT ZN\_FING 265 287 C2H2-TYPE.  
 FT ZN\_FING 293 315 C2H2-TYPE.  
 FT ZN\_FING 321 343 C2H2-TYPE.  
 SQ SEQUENCE 348 AA; 40297 MW; E0941AD23EC8670 CRC64;  
 Query Match 71.2%; Score 99; DB 1; Length 348;  
 Best Local Similarity 68.0%; Pred. No. 2.2e-07;  
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 YQCEICGKSPKSDKSNLTTHRIHTG 25  
 Db 153 YCESGCGKTFGRKDNLTQHKRIHTG 177  
 RESULT 11  
 ZP58\_MOUSE STANDARD; PRT; 169 AA.  
 AC P16372;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Zinc finger protein 58 (Zfp-58) (Zinc finger protein Mfg-1)  
 DE (Fragment).  
 GN ZP58 OR MFG-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Skeletal muscle;  
 RA MEDLINE=90083278; PubMed=2512579;  
 RA Passaniti C., Felsani A., Carneio M., Amati P.;  
 RA "Mouse genes coding for 'zinc-finger'-containing proteins:  
 RA characterization and expression in differentiated cells.";  
 RA Proc. Natl. Acad. Sci. U.S.A. 86:9417-9421(1989)  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:9417-9421(1989)  
 CC -1- FUNCTION: MAY HAVE A ROLE DURING DIFFERENTIATION PROCESSES.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- TISSUE SPECIFICITY: LIVER; TESTIS; AND AT CONSIDERABLY LOWER  
 CC LEVELS IN BRAIN; SPLEEN; HEART.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS POSITIVELY REGULATED UPON  
 CC DIFFERENTIATION, AND IS NOT RELATED TO THE CELL CYCLE.  
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M28513; AAA95531.1; -  
 CC PIR: A39240; A39240.  
 CC HSSP: P25490; 1UBD.  
 CC MGD: MG1:99205; Zfp58.  
 CC InterPro: IPR007087; Znf\_C2H2.  
 CC InterPro: IPR007086; Znf\_C2H2\_sub.  
 CC Pfam: PF00096; zf-C2H2; 6.  
 CC PRINTS: PR00048; ZINC\_FINGER.  
 CC SMART: SM00355; Znf\_C2H2; 6.  
 CC PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 5.  
 CC PROSITE: PS50157; ZINC\_FINGER\_C2H2\_2; 6.  
 CC Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.  
 FT ZN\_FING 1 1 C2H2-TYPE.  
 FT ZN\_FING 3 30 C2H2-TYPE.  
 FT ZN\_FING 36 58 C2H2-TYPE.  
 FT ZN\_FING 64 86 C2H2-TYPE.  
 FT ZN\_FING 92 114 C2H2-TYPE.  
 FT ZN\_FING 120 142 C2H2-TYPE.  
 FT ZN\_FING 148 169 C2H2-TYPE.  
 FT NON\_TER 169 169  
 SQ SEQUENCE 169 AA; 19585 MW; 7F55256B1406CDA3 CRC64;  
 Query Match 70.5%; Score 98; DB 1; Length 169;  
 Best Local Similarity 68.0%; Pred. No. 1.4e-07;  
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 YQCEICGKSPKSDKSNLTTHRIHTG 25  
 Db 92 YXCEGCGKAFSTSNLSERKRIHTG 116  
 RESULT 12  
 ZP38\_HUMAN STANDARD; PRT; 473 AA.  
 AC Q91SAG; Q9H0B5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Zinc finger protein 38 homolog (Zfp-38) (NY-REN-21 antigen).  
 GN ZFP38.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA MEDLINE=21154917; PubMed=11230166;  
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Giesel S.,  
 RA Ansborge W., Boecker H., Baerachsch S., Blum H.,  
 RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,  
 RA Mewes H.-W., Oltmannseder B., Obermayer B., Tampe J., Heubner D.,  
 RA Mamutic R., Korn B., Klein M., Poustka A.;  
 RA "Towards a catalog of human genes and proteins: sequencing and  
 RA analysis of 500 novel complete protein coding human cDNAs";  
 RA Genome Res. 11:422-435(2001).  
 RL [2]  
 RP SEQUENCE OF 62-473 FROM N.A.  
 RC TISSUE=renal cell carcinoma;  
 RA MEDLINE=99438124; PubMed=10508479;  
 RA Scallan M.J., Gordan J.D., Williamson B., Stockert E., Bander N.H.,  
 RA Joneneel C.V., Gure A.O., Jager D., Jager E., Knuth A., Chen Y.-T.,  
 RA Old L.T.;  
 RA "Antigens recognized by autologous antibody in patients with renal-  
 RA cell carcinoma.";  
 RT Int. J. Cancer 83:456-464(1999).  
 CC -----

```

CC - FUNCTION: STRONG TRANSCRIPTIONAL ACTIVATOR (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: Nuclear (By similarity).
CC - SIMILARITY: BELONGS TO THE KRUPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC - SIMILARITY: Contains 1 SCAN box domain.
CC - SIMILARITY: Contains 7 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, A136865; CAB66799.1; -.
DR EMBL, AF15100; AAD2866.1; -.
DR HSSP, P08153; 1ZFD.
DR GO, GO:0005634; C:nucleus; NAS.
DR GO, GO:0016563; F:transcriptional activator activity; NAS.
DR GO, GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro, IPR003309; Trp-SCAN.
DR InterPro, IPR007087; ZnF_C2H2.
DR Pfam, PF02023; SCAN; 1.
DR Pfam, PF00096; zf_C2H2; 7.
DR ProDom, PD000003; ZnF_C2H2; 7.
DR SMART, SM00355; ZNF_C2H2; 7.
DR SMART, SM00804; SCAN_BOX; 1.
DR PROSITE, PS00028; ZINC_FINGER_C2H2_1; 6.
DR PROSITE, PS50157; ZINC_FINGER_C2H2_2; 7.
DR Zinc-finger, Metal-binding; DNA-binding; Nuclear protein; Repeat;
DR Transcription regulation; Activator.
DR DOMAIN 45 127
FT ZN_FING 277 289 C2H2-TYPE 1.
FT ZN_FING 305 327 C2H2-TYPE 2.
FT ZN_FING 333 354 C2H2-TYPE 3.
FT ZN_FING 360 382 C2H2-TYPE 4.
FT ZN_FING 388 410 C2H2-TYPE 5.
FT ZN_FING 416 438 C2H2-TYPE 6.
FT ZN_FING 444 466 C2H2-TYPE 7.
SQ SEQUENCE 473 AA; 53658 MW; D9A2664B14B96F CRC64;

Query Match 70.5%; Score 98; DB 1; Length 473;
Best Local Similarity 68.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 YOCBICGKSFSDKSNLTRLRHTG 25
Db 360 YQCKDCGKAFSGKSGSLIRYRHTG 384

RESULT 13
ID 2256 HUMAN STANDARD; PRT; 474 AA.
AC Q9Y2F7; Q9Y2F7.1; 40. Created.
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 256 (Bone marrow zinc finger 3) (BMZF-3).
GN ZNF256 OR BMZF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow; PubMed=10585455;
RX MEDLINE=20054457;
RA Shi S.-L., Zhou D., Fu G., Mao W., Chen S.-J., Yu L., Chen Z.,
RT "Molecular cloning of six novel Kruppel-like zinc finger genes from
hematopoietic cells and identification of a novel transregulatory

```

```

RT domain KRNB."
RL J. Biol. Chem. 274:35741-35748 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schlier G.D.,
RA Altschul S.F., Zeeberg B., Bietow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang T., Hsieh F.,
RA Diatchenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo A.F., Cassavant T.L., Schaeetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzyzanski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Maria M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC - FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
CC - SUBCELLULAR LOCATION: Nuclear (Probable).
CC - SIMILARITY: BELONGS TO THE KRUPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF067165; AAD32449.1; -.
DR EMBL, BC001438; AAF01438.1; -.
DR HSSP, P08045; 1ZNF.
DR Genew; HGNC:13049; ZNF256.
DR MIM; 606956; -.
DR GO, GO:0003700; F:transcription factor activity; TAS.
DR GO, GO:0007275; P:development; TAS.
DR InterPro, IPR007087; ZnF_C2H2.
DR Pfam, PF00096; zf_C2H2; 14.
DR ProDom, PD000003; ZnF_C2H2; 12.
DR SMART, SM00355; ZNF_C2H2; 14.
DR PROSITE, PS00028; ZINC_FINGER_C2H2_1; 14.
DR PROSITE, PS50157; ZINC_FINGER_C2H2_2; 14.
DR Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
DR Nuclear protein; Repeat.
KV Nuclear protein; Repeat.
FT ZN_FING 86 108 C2H2-TYPE.
FT ZN_FING 114 136 C2H2-TYPE.
FT ZN_FING 142 164 C2H2-TYPE.
FT ZN_FING 170 192 C2H2-TYPE.
FT ZN_FING 198 220 C2H2-TYPE.
FT ZN_FING 226 248 C2H2-TYPE.
FT ZN_FING 254 276 C2H2-TYPE.
FT ZN_FING 282 304 C2H2-TYPE.
FT ZN_FING 310 332 C2H2-TYPE.
FT ZN_FING 338 360 C2H2-TYPE.
FT ZN_FING 366 388 C2H2-TYPE.
FT ZN_FING 394 416 C2H2-TYPE.
FT ZN_FING 422 444 C2H2-TYPE.
FT ZN_FING 444 472 C2H2-TYPE.
FT ZN_FING 450 472 C2H2-TYPE.
SQ SEQUENCE 474 AA; 54872 MW; 2C524C25743A9014 CRC64;

Query Match 70.5%; Score 98; DB 1; Length 474;
Best Local Similarity 68.0%; Pred. No. 4.3e-07;

```

Matches 17; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 YQCEICGKSPGSKSLTRHRIHTG 25  
 Db 442 YQCEICGKSPGSKSLTRHRIHTG 446

RESULT 14  
 ZF38\_MOUSE  
 ID\_ZF38\_MOUSE STANDARD; PRT; 555 AA.

AC 007231;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DE Zinc finger protein 38 (Zfp-38) (CfPINS1) (Transcription factor RU49).  
 GN ZNF38 OR ZIPF01 OR ZFP38 OR ZFP-38.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BTBRF; TISSUE=Spermatocyte;  
 RA MEDLINE=93012481; PubMed=1397691;  
 RX Noce T., Fujiwara Y., Sezaki M., Fujimoto H., Higashinakagawa T.,  
 RT "Expression of a mouse zinc finger protein gene in both spermatocytes  
 and oocytes during meiosis."  
 RL Dev. Biol. 153:356-367(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RA MEDLINE=93183757; PubMed=1284028;  
 RX Chowdhury K.;  
 RT "The ubiquitous transactivator Zfp-38 is upregulated during  
 spermatogenesis with differential transcription."  
 RL Mech. Dev. 39:129-142(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA MEDLINE=96189345; PubMed=8625807;  
 RX Yang X.W., Zhong R., Heintz N.;  
 RT "Granule cell specification in the developing mouse brain as defined  
 by expression of the zinc finger transcription factor RU49."  
 RL Development 122:555-566(1996).  
 CC -1- FUNCTION: STRONG TRANSCRIPTIONAL ACTIVATOR. ASSOCIATED WITH  
 MEIOSIS IN BOTH MALE AND FEMALE GAMETOGENESIS. MAY HAVE DIFFERENT  
 FUNCTIONS IN SOMATIC CELLS.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE SPERMATOCYTES AND  
 SPERMATIDS OF ADULT TESTES. IT IS ALSO PRESENT AT LOWER LEVELS IN  
 THE OVARY, BRAIN, SPLEEN, EMBRYO AND FETUS.  
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED BETWEEN 2 AND 3 WEEKS AFTER  
 BIRTH, IN PARALLEL WITH THE ONSET AND PROGRESSION OF MEIOSIS. IT  
 IS EXPRESSED DURING OOOGENESIS FROM THE PACHTENE STAGE OF MEIOTIC  
 PROPHASE THROUGH TO POSTMEIOTIC CELLS.  
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
 FINGER PROTEINS.  
 CC -1- SIMILARITY: Contains 1 SCAN box domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D10630; BA001480.1; -;  
 DR EMBL; X63747; CA445280.1; -;  
 DR EMBL; U41671; AAB03786.1; -;  
 DR FPI; A56560; A56560.  
 DR HSSP; P08045; 1ZNF.  
 DR MGD; MGI:99182; Zfp38l.

DR InterPro; IPR003309; Treg\_SCAN.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR Pfam; PF002023; SCAN; 1.  
 DR Pfam; PF00096; ZF-C2H2; 7.  
 DR ProDom; PD000003; Znf\_C2H2; 6.  
 DR SMART; SM00431; LER; 1.  
 DR SMART; SM00355; ZNF\_C2H2; 7.  
 DR PROSITE; PS50804; SCAN\_BOX; 1.  
 DR PROSITE; PS00026; ZINC\_FINGER\_C2H2\_1; 6.  
 DR PROSITE; PS01571; ZINC\_FINGER\_C2H2\_2; 7.  
 DR Zinc-finger, Metal-binding; DNA-binding; Nuclear protein; Repeat;  
 KW Developmental protein; Spermatogenesis; Transcription regulation;  
 FT DOMAIN 18 134 3 X 39 AA APPROXIMATE TANDEM REPEATS.  
 FT REPEAT 18 56  
 FT REPEAT 17 95 1-1.  
 FT REPEAT 96 134 1-2.  
 FT REPEAT 122 204 1-3.  
 FT DOMAIN 122 204 SCAN\_BOX.  
 FT ZN\_FING 359 381 C2H2-TYPE.  
 FT ZN\_FING 387 409 C2H2-TYPE.  
 FT ZN\_FING 415 436 C2H2-TYPE.  
 FT ZN\_FING 442 464 C2H2-TYPE.  
 FT ZN\_FING 470 492 C2H2-TYPE.  
 FT ZN\_FING 498 520 C2H2-TYPE.  
 FT ZN\_FING 526 548 C2H2-TYPE.  
 FT CONFLICT 193 193 T -> A (IN REF. 2 AND 3).  
 FT CONFLICT 216 216 N -> S (IN REF. 3).  
 FT CONFLICT 234 234 S -> P (IN REF. 3).  
 FT CONFLICT 334 347 ASLDTGSRGAEP -> PLFKTVPRGGA (IN REF. 3).  
 FT CONFLICT 401 402 NL -> KV (IN REF. 3).  
 FT CONFLICT 478 478 S -> E (IN REF. 3).  
 FT CONFLICT 495 495 E -> K (IN REF. 2).  
 FT CONFLICT 507 507 F -> L (IN REF. 2).  
 FT CONFLICT 510 510 S -> N (IN REF. 2).  
 SQ SEQUENCE 555 AA; 63042 MW; CAEBBF0A149C6C6 CRC64;

Query Match Best Local Similarity 70.5%; Score 98; DB 1; Length 555;  
 Match 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSPGSKSLTRHRIHTG 25  
 Db 442 YQCEICGKSPGSKSLTRHRIHTG 466

RESULT 15  
 GLAS\_DROME  
 ID\_GLAS\_DROME STANDARD; PRT; 604 AA.  
 AC P13360;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Glab protein.  
 GN GL.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydriidae; Drosophilidae; Drosophila.  
 CX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=89365138; PubMed=2770860;  
 RA Moses K., Ellis M.C., Rubin G.M.;  
 RT "The glab gene encodes a zinc-finger protein required by Drosophila  
 RT photoreceptor cells."  
 RL Nature 340:531-536(1989).  
 CC -1- FUNCTION: GLAB IS PROBABLY A TRANSCRIPTION FACTOR REQUIRED FOR  
 GENE EXPRESSION SPECIFIC TO PHOTORECEPTOR CELLS.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (potential).  
 CC -1- SIMILARITY: Contains 5 C2H2-type zinc fingers.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL/ X15400; CAA33450.1; -.  
 DR PIR: S05447; S05447.  
 DR HSSP: P08046; 1A1I.  
 DR TRANSFAC: T00328; -.  
 DR FLYbase; FBgn0004618; g1.  
 DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0007459; P:photoreceptor commitment (sensu Drosophila); IMP.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR Pfam; PF00096; zf-C2H2; 5.  
 DR ProDom; PD000003; Znf\_C2H2; 2.  
 DR SMART; SM00355; Znf\_C2H2; 5.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 5.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 5.  
 KW Transcription regulation; zinc-finger; Metal-binding; DNA-binding;  
 KW Repeat; Vision; Nuclear protein.  
 FT ZN\_FING 437 459 C2H2-TYPE 1.  
 FT ZN\_FING 465 487 C2H2-TYPE 2.  
 FT ZN\_FING 493 515 C2H2-TYPE 3.  
 FT ZN\_FING 521 543 C2H2-TYPE 4.  
 FT ZN\_FING 549 571 C2H2-TYPE 5.  
 FT VARIANT 60 60 E -> O.  
 FT VARIANT 361 362 PM -> SL.  
 FT VARIANT 377 377 S -> T.  
 SQ SEQUENCE 604 AA; 62528 MW; 82B947C78366C071 CRC64;  
 QY Query Match 70.5%; Score 98; DB 1; Length 604;  
 DB Best Local Similarity 68.0%; Pred. No. 5.5e-07;  
 1 YQCEICGKSFSDKXNLTRHLRIHTG 25  
 521 YRCSCKKSFSDSTLTKLRLIHSG 545  
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Search completed: January 6, 2004, 16:11:03  
 Job time : 12 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 16:09:51 : Search time 22 Seconds  
(without alignments)  
109.283 Million cell updates/sec

Title: US-09-774-940A-1

Sequence: 1 YQCEICGKSFSDKSNLTRHLRHTG 25

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	106	76.3	242 2	S30238
2	102	73.4	106 2	T34560
3	101	72.7	485 2	S52411
4	100	71.9	247 2	S47070
5	100	71.9	595 2	G02075
6	100	71.9	1191 2	S35305
7	99	71.2	348 2	I38599
8	98	70.5	169 2	A39240
9	98	70.5	555 2	A56560
10	98	70.5	604 2	S05447
11	97	69.8	319 2	T46469
12	96	69.1	88 2	D43284
13	96	69.1	325 2	T13816
14	96	69.1	386 2	T12527
15	96	69.1	424 2	S15917
16	96	69.1	469 2	I38600
17	96	69.1	594 2	T12488
18	96	69.1	803 2	S26823
19	95	68.3	149 2	M28942
20	95	68.3	378 2	S33994
21	95	68.3	405 2	T46663
22	95	68.3	427 2	A35659
23	95	68.3	428 2	A33891
24	95	68.3	651 2	B33891
25	94	67.6	56 2	I37958
26	94	67.6	223 2	S33991
27	94	67.6	243 2	T13225
28	94	67.6	292 2	S43826
29	94	67.6	457 2	C57785

30	94	67.6	485	2	A40751	finger protein MZF
31	94	67.6	701	2	T14757	hypothetical prote
32	93	66.9	347	2	S00549	developmental cont
33	93	66.9	488	2	S47072	finger protein HZF
34	93	66.9	540	2	T14748	hypothetical prote
35	93	66.9	542	2	A54661	zinc finger protei
36	93	66.9	580	2	A37107	spermatogenesis pr
37	93	66.9	589	2	I38598	zinc finger protei
38	93	66.9	614	2	TJ0500	zinc finger protei
39	93	66.9	725	2	T42868	hypothetical prote
40	92	66.2	52	2	I37971	zinc finger protei
41	92	66.2	200	2	T26715	hypothetical prote
42	92	66.2	295	2	A46017	zinc finger protei
43	92	66.2	428	2	S03677	finger protein (cl
44	92	66.2	636	2	I48689	Gene NK10 protein
45	92	66.2	710	2	I48668	zinc finger protei

## ALIGNMENTS

## RESULT 1

S30238  
finger protein ZNF37A - human (fragment)

C/Species: Homo sapiens (man)  
C/Date: 22-Nov-1993 #sequence, revision 26-May-1995 #ext\_change 08-Dec-2000

C/Accession: S30238; I37961; S10410  
R/Turnaciliffe, A.; Liu, L.; Moore, J.K.; Leverisha, M.A.; Jackson, M.S.; Papi, L.; Fergu

Nucleic Acids Res. 21, 1409-1417, 1993  
A/Title: Duplicated KOX zinc finger gene clusters flank the centromere of human chromos

A/Reference number: S30238; MIMD:33219119; PMID:8464732  
A/Accession: S30238

A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: mRNA

A/Residues: 1-242 <TUN>  
A/Cross-references: EMBL:X69115; NID:G288424; PID:CAA48868.1; PID:G93084

A/Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1992  
R/TJensen, H.J.

New Biol. 2, 363-374, 1990  
A/Title: Multiple genes encoding zinc finger domains are expressed in human T cells.

A/Reference number: I37961; MIMD:91145339; PMID:2288909  
A/Accession: I37961

A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA

A/Residues: 8-63 <RES>  
A/Cross-references: EMBL:X52352; NID:G34154; PID:CAA36578.1; PID:G93084

A/Note: zinc finger protein ZFP-36; LIM metal-binding repeat homology  
C/Keywords: DNA binding; zinc finger

Query Match  
Best local similarity 76.3%; Score 106; DB 2; Length 242;  
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSFSDKSNLTRHLRHTG 25

Db 92 YECNECGKSFSEKSLTKHLRHTG 116

## RESULT 2

T34560  
hypothetical protein DKFZp434U0650.1 - human (fragment)

C/Species: Homo sapiens (man)  
C/Date: 29-Oct-1999 #sequence, revision 29-Oct-1999 #ext\_change 29-Oct-1999

C/Accession: T34560  
R/Pouskka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, October 1999  
A/Reference number: Z21540

A/Accession: T34560  
A/Status: preliminary

A/Molecule type: mRNA  
A/Residues: 1-106 <POU>

A/Cross-references: EMBL:AL122078  
A/Experimental source: adult testis; clone DKFZp434U0650



C/Genetics:  
A/Gene: GDB:ZNF134  
A/Cross-references: GDB:137033  
A/Map position: 19q13.4-19q13.4

Query Match 71.2%; Score 99; DB 2; Length 348;  
Best Local Similarity 68.0%; Pred. No. 1.3e-06;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDKSNLTRHLRIHTG 25  
DB 153 YQCEGCKAFSTSNLSEHKRIHTG 177

RESULT 8  
A39240  
finger protein mfg1 - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 04-Oct-1991 #sequence\_revision 04-Oct-1991 #text\_change 05-Nov-1999  
C/Accession: A39240  
R/Passananti, C.; Felisani, A.; Carnuso, M.; Amati, P.  
Proc. Natl. Acad. Sci. U.S.A. 86, 9417-9421, 1989  
A/Title: Mouse gene coding for "zinc-finger"-containing proteins: characterization and  
A/Reference number: A39240; MUID:90083278; PMID:2512579  
A/Accession: A39240  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-169 <PAS>  
A/Cross-references: GB:M28513; NID:9199136; PIDN:AAA39531.1; PID:9554203  
C/Keywords: DNA binding; zinc finger

Query Match 70.5%; Score 98; DB 2; Length 169;  
Best Local Similarity 68.0%; Pred. No. 9.4e-07;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDKSNLTRHLRIHTG 25  
DB 92 YQCEGCKAFSTSNLSEHKRIHTG 116

RESULT 9  
A56560  
zinc finger protein/transactivator Zfp-38 - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 21-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 01-Dec-2000  
C/Accession: A56560; C48827  
R/Choudhury, K.; Goulding, M.; Walther, C.; Imai, K.; Fickenscher, H.  
Mech. Dev. 39, 129-142, 1992  
A/Title: The ubiquitous transactivator Zfp-38 is upregulated during spermatogenesis with  
A/Reference number: A56560; MUID:93183757; PMID:1284028  
A/Accession: A56560  
A/Molecule type: mRNA  
A/Residues: 1-555 <CHO>  
A/Cross-references: GB:X63747; NID:955476; PIDN:CAA5280.1; PID:955477  
A/Experimental source: embryo  
A/Note: sequence extracted from NCBI backbone (NCBIN:126596, NCBI:P:126599)  
R/Noce, T.; Fujiwara, Y.; Sezaki, M.; Fujimoto, H.; Higashinakagawa, T.  
Dev. Biol. 153, 356-367, 1992  
A/Title: Expression of a mouse zinc finger protein gene in both spermatocytes and oocytes  
A/Reference number: A48827; MUID:93012481; PMID:1397691  
A/Accession: C48827  
A/Molecule type: mRNA

A/Residues: 1-192, 'T', 194-494, 'E', 496-506, 'F', 508-509, 'S', 511-555 <NOC>  
A/Cross-references: EMBL:D10630; NID:9220640; PIDN:BA01480.1; PID:9220641  
A/Experimental source: spermatogenic cells; clone CTF151  
A/Note: sequence extracted from NCBI backbone (NCBI:P:114773)  
C/Genetics:  
A/Gene: Zfp-38  
A/Map position: 5  
C/Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology  
C/Keywords: spermatogenesis; transcription regulation; zinc finger

Query Match 70.5%; Score 98; DB 2; Length 555;

Best Local Similarity 68.0%; Pred. No. 2.7e-06;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDKSNLTRHLRIHTG 25  
DB 442 YQCEGCKAFSTSNLSEHKRIHTG 466

## RESULT 10

S05447  
finger protein glass - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 21-Jul-2000  
C/Accession: S05447  
R/Moses, K.; Ellis, M.C.; Rubin, G.M.  
Nature 340, 531-536, 1989  
A/Title: The glass gene encodes a zinc-finger protein required by Drosophila photorecep  
A/Reference number: S05447; MUID:89355138; PMID:2770860  
A/Accession: S05447  
A/Molecule type: DNA  
A/Residues: 1-604 <MOS>  
A/Cross-references: EMBL:X15400; NID:98015; PIDN:CAA33450.1; PID:98016  
A/Note: 60-Gln, 361-Ser, 362-Leu, and 377-Thr were also found  
C/Genetics:  
A/Gene: glass  
A/Cross-references: FlyBase:FBgn0004618  
A/Map position: 3R 91A1-2  
A/Insertions: 16/1; 59/3; 557/2  
C/Keywords: DNA binding; nucleus; transcription factor; zinc finger  
F/439-571/Region: zinc fingers

Query Match 70.5%; Score 98; DB 2; Length 604;  
Best Local Similarity 68.0%; Pred. No. 2.9e-06;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDKSNLTRHLRIHTG 25  
DB 521 YQCEGCKAFSTSNLSEHKRIHTG 545

## RESULT 11

T46469  
hypothetical protein DKFZp434G1930.1 - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C/Accession: T46469  
R/Blum, H.; Bauerbachs, S.; Mewes, H.W.; Gassenhuber, U.; Wiemann, S.  
Submitted to the Protein Sequence Database, January 2000  
A/Reference number: Z23034  
A/Accession: T46469  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-319 <AA>  
A/Cross-references: EMBL:AL137483  
A/Experimental source: adult testis; clone DKFZp434G1930  
C/Genetics:  
A/Note: DKFZp434G1930.1

Query Match 69.8%; Score 97; DB 2; Length 319;  
Best Local Similarity 72.0%; Pred. No. 2.3e-06;  
Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDKSNLTRHLRIHTG 25  
DB 205 YQCEGCKAFSTSNLSEHKRIHTG 229

## RESULT 12

D43284  
finger protein ZNF50 - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998  
C/Accession: D43284

R.Lichter, P.; Bray, P.; Ried, T.; Dawid, I.B.; Ward, D.C.  
 Genomics 13, 999-1007, 1992  
 A>Title: Clustering of C2-H2 zinc finger motif sequences within telomeric and fragile si  
 A/Reference number: A43284; MUID:92372070; PMID:1505991  
 A/Accession: D43284  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-88 <LIC>  
 A/Cross-references: GB:M8360; NID:G340457; PID:G340458  
 A/Note: sequence extracted from NCBI backbone (NCBIN:111638, NCBIPI:111639)  
 C/Keywords: zinc finger

Query Match 69.1%; Score 96; DB 2; Length 89;  
 Best Local Similarity 60.0%; Pred. No. 1e-06;  
 Matches 15; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDKNLTTRHRIHTG 25  
 DB 11 YECDAKGAFQSGSTHLTQHORHTG 35

# RESULT 13

1386:6  
 zinc finger protein ZNF139 - human (fragment)  
 N/Alternate names: finger protein kox18  
 C/Species: Homo sapiens (man)  
 C/Date: 23-Feb-1996 #sequence revision 23-Feb-1996 #text\_change 01-Dec-2000  
 C/Accession: 138616; 137957; S10406  
 R/Tommerup, N.; Vissing, H.  
 Genomics 27, 259-264, 1995  
 A>Title: Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs identifi  
 A/Reference number: A57785; MUID:9604430; PMID:7557990  
 A/Accession: 138616  
 A/Molecule type: mRNA  
 A/Residues: 1-325 <RES>  
 A/Cross-references: EMBL:U09848; NID:9495567; PIDN:AA50264.1; PID:9495568  
 R/Thiesen, H.J.  
 New Biol. 2, 363-374, 1990  
 A>Title: Multiple genes encoding zinc finger domains are expressed in human T cells.  
 A/Reference number: 137949; MUID:91143339; PMID:2288909  
 A/Accession: 137957  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 167-222 <RE2>  
 A/Cross-references: EMBL:X52349; NID:G34150; PIDN:CA36575.1; PID:G330080  
 C/Genetics:  
 A/Gene: GDB:ZNF139  
 A/Cross-references: GDB:137044  
 A/Map position: 7q21.3-7q22.1  
 C/Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology  
 C/Keywords: DNA binding; zinc finger

Query Match 69.1%; Score 96; DB 2; Length 325;  
 Best Local Similarity 64.0%; Pred. No. 3.3e-06;  
 Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDKNLTTRHRIHTG 25  
 DB 223 YECNECGKAFQSSDLTKGRIHTG 247

# RESULT 14

112527  
 hypothetical protein DKFZp434N043.1 - human (fragment)  
 C/Species: Homo sapiens (man)  
 C/Date: 23-Jul-1999 #sequence revision 23-Jul-1999 #text\_change 01-Dec-2000  
 C/Accession: 112527  
 R/Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, U.; Wiemann, S.  
 submitted to the Protein Sequence Database, June 1999  
 A/Reference number: Z17524  
 A/Accession: 112527  
 A/Status: preliminary  
 A/Molecule type: mRNA

A/Residues: 1-386 <MAN>  
 A/Cross-references: EMBL:AL080143  
 A/Experimental source: adult testis; clone DKFZp434N043  
 C/Genetics:  
 A/Note: DKFZp434N043.1  
 C/Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 69.1%; Score 96; DB 2; Length 386;  
 Best Local Similarity 64.0%; Pred. No. 3.8e-06;  
 Matches 16; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDKNLTTRHRIHTG 25  
 DB 301 YECGCGKAFQKANTLTQHORHTG 325

# RESULT 15

S15917  
 developmental control protein mkrl - mouse  
 N/Alternate names: finger protein zfp-1  
 C/Species: Mus musculus (house mouse)  
 C/Date: 04-Dec-1992 #sequence revision 04-Dec-1992 #text\_change 01-Dec-2000  
 C/Accession: S15917; A26595; S06773  
 R/Chowdhury, K.; Dietrich, S.; Balling, R.; Guenet, J.L.; Gruss, P.  
 Nucleic Acids Res. 17, 10427-10438, 1989  
 A>Title: Structure, expression and chromosomal localization of Zfp-1, a murine zinc fi  
 A/Reference number: S15917; MUID:90098878; PMID:2574853  
 A/Accession: S15917  
 A/Molecule type: mRNA  
 A/Residues: 1-424 <CHO>  
 A/Cross-references: EMBL:X16493; NID:G55482; PIDN:CA34510.1; PID:G55483  
 R/Chowdhury, K.; Deusch, U.; Gruss, P.  
 Cell 48, 771-778, 1987  
 A>Title: A multigene family encoding several "finger" structures is present and differ  
 A/Reference number: A30892; MUID:87131089; PMID:3815523  
 A/Accession: A26595  
 A/Molecule type: DNA  
 A/Residues: 'AAG', 54-424 <CH2>  
 A/Cross-references: GB:M15708; NID:G193349; PIDN:AA37639.1; PID:G387162  
 C/Genetics:  
 A/Gene: zfp-1  
 C/Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology  
 C/Keywords: DNA binding; zinc finger

Query Match 69.1%; Score 96; DB 2; Length 424;  
 Best Local Similarity 60.0%; Pred. No. 4.2e-06;  
 Matches 15; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDKNLTTRHRIHTG 25  
 DB 197 YECVCGKTFSHKANTLTQHORHTG 221

Search completed: January 6, 2004, 16:13:00  
 Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Inc.

OM protein - protein search, using sw model

Run on: January 6, 2004, 16:11:57; Search time 33 Seconds

(without alignments)  
152.001 Million cell updates/sec

Title: US-09-774-940A-1

Perfect score: 139  
Sequence: 1 YQCEICKGSFSDKSNLTRHLRIHTG 25

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCTI\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTIS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*

Print. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	139	100.0	25	10	US-09-774-940A-1
2	108	77.7	29	9	US-09-851-271A-9
3	108	77.7	29	11	US-09-995-973-9
4	108	77.7	29	11	US-09-996-484-9
5	108	77.7	29	12	US-10-192-078-23
6	106	76.3	561	12	US-10-104-047-2805
7	104	74.8	362	15	US-10-043-487-390
8	104	74.8	525	9	US-09-864-761-33653
9	104	74.8	576	12	US-10-094-749-2365
10	103	74.1	409	12	US-10-104-047-2663
11	103	74.1	475	12	US-10-104-047-2470
12	102	73.4	25	15	US-10-198-677-78
13	102	73.4	530	12	US-10-104-047-7434
14	100	71.9	242	12	US-10-029-386-33120
15	100	71.9	295	12	US-10-094-749-1859

16	100	71.9	504	12	US-10-104-047-2053	Sequence 2053, Ap
17	99	71.2	378	12	US-10-314-669-2	Sequence 2, Appl
18	99	71.2	413	9	US-09-864-761-37010	Sequence 37010, A
19	98	70.5	107	12	US-10-314-669-23	Sequence 23, Appl
20	98	70.5	115	10	US-09-925-300-1555	Sequence 1555, Ap
21	98	70.5	125	9	US-09-764-853-815	Sequence 815, Ap
22	98	70.5	209	12	US-10-314-669-10	Sequence 10, Appl
23	98	70.5	235	9	US-09-764-853-633	Sequence 633, Ap
24	98	70.5	433	15	US-10-171-311-232	Sequence 232, Ap
25	98	70.5	446	12	US-10-094-749-2105	Sequence 2105, Ap
26	98	70.5	1357	12	US-10-295-027-1199	Sequence 1199, Ap
27	97	69.8	88	9	US-09-851-271A-2	Sequence 2, Appl
28	97	69.8	111	10	US-09-764-864-1502	Sequence 1502, Ap
29	97	69.8	111	11	US-09-764-891-4219	Sequence 4219, Ap
30	97	69.8	127	12	US-10-029-386-30375	Sequence 30375, A
31	97	69.8	134	10	US-09-764-864-1080	Sequence 1080, Ap
32	97	69.8	196	9	US-09-864-761-48174	Sequence 48174, A
33	97	69.8	353	12	US-10-029-386-32920	Sequence 32920, A
34	97	69.8	402	12	US-10-094-749-2063	Sequence 2063, Ap
35	97	69.8	406	12	US-10-094-749-2371	Sequence 2371, Ap
36	97	69.8	545	12	US-10-029-386-33143	Sequence 33143, A
37	97	69.8	555	12	US-10-094-749-2231	Sequence 2231, Ap
38	97	69.8	558	9	US-09-864-761-47724	Sequence 47724, A
39	97	69.8	592	12	US-10-104-047-3366	Sequence 3366, Ap
40	97	69.8	601	12	US-10-029-386-32907	Sequence 32907, A
41	97	69.8	780	12	US-10-094-749-2634	Sequence 2634, Ap
42	97	69.8	811	12	US-10-144-194A-26	Sequence 2634, Ap
43	97	69.8	811	12	US-10-144-194A-26	Sequence 2634, Ap
44	97	69.8	816	12	US-10-104-047-2341	Sequence 2341, Ap
45	96	69.1	156	12	US-10-029-386-30943	Sequence 30943, A

#### ALIGNMENTS

RESULT 1  
US-09-774-940A-1  
; Sequence 1, Application US/09774940A  
; Patent No. US20020155576A1  
; GENERAL INFORMATION:  
; APPLICANT: MILLS, STRANLEY L.  
; APPLICANT: MILLS, JACQUELINE L.  
; TITLE OF INVENTION: METAL-CHELATED NUCLEIC ACID BINDING PEPTIDES FOR IN VIVO DETECTION  
; FILE REFERENCE: 5157,002  
; CURRENT APPLICATION NUMBER: US/09/774,940A  
; CURRENT FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: US 09/021,085  
; PRIOR FILING DATE: 1998-02-10  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: ARTIFICIAL SEQUENCE  
; FEATURE:  
; OTHER INFORMATION: CYS-CYS-HIS-HIS ZINC FINGER  
US-09-774-940A-1

Query Match 100.0%; Score 139; DB 10;  
Best Local Similarity 100.0%; Pred. No. 5.9e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQCEICKGSFSDKSNLTRHLRIHTG 25  
Db 1 YQCEICKGSFSDKSNLTRHLRIHTG 25

RESULT 2  
US-09-851-271A-9  
; Sequence 9, Application US/09851271A  
; Patent No. US2002064824A1  
; GENERAL INFORMATION:

```
; APPLICANT: Gendag Limited
; TITLE OF INVENTION: Screening System
; FILE REFERENCE: 674538-2003
; CURRENT APPLICATION NUMBER: US/09/851.271A
; CURRENT FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: PCT/GB99/03730
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: GB9824544.2
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: ZN.FING
; LOCATION: (1)..(29)
; OTHER INFORMATION: zinc finger consensus structure
US-09-851-271A-9
```

```
Query Match          77.7%; Score 108; DB 9; Length 29;
Best Local Similarity 76.0%; Pred. No. 1.7e-08;
Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 1 YQCEICGKSFSDKSNLTRHLRIHTG 25
Db 2 YKCEGCKAFSQKSNLTRHQRHTG 26
```

```
RESULT 3
US-09-995-973-9
; Sequence 9, Application US/09995973
; Publication No. US20030024006A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: GENE SWITCHES
; FILE REFERENCE: 8325-2003 / G7-US1
; CURRENT APPLICATION NUMBER: US/09/995.973
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
US-09-995-973-9
```

```
Query Match          77.7%; Score 108; DB 11; Length 29;
Best Local Similarity 76.0%; Pred. No. 1.7e-08;
Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 1 YQCEICGKSFSDKSNLTRHLRIHTG 25
Db 2 YKCEGCKAFSQKSNLTRHQRHTG 26
```

```
RESULT 4
US-09-996-484-9
; Sequence 9, Application US/09996484
; Publication No. US20030092010A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: MOLECULAR SWITCHES
; FILE REFERENCE: 8325-2004 / G8-US1
; CURRENT APPLICATION NUMBER: US/09/996.484
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 64
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
US-09-996-484-9
```

```
Query Match          77.7%; Score 108; DB 11; Length 29;
Best Local Similarity 76.0%; Pred. No. 1.7e-08;
Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 1 YQCEICGKSFSDKSNLTRHLRIHTG 25
Db 2 YKCEGCKAFSQKSNLTRHQRHTG 26
```

```
RESULT 5
US-10-192-078-23
; Sequence 23, Application US/10192078
; Publication No. US20030188331A1
; GENERAL INFORMATION:
; APPLICANT: Choo, Yen
; APPLICANT: Ullman, Christopher Graeme
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Sanchez, Juan Pablo
; TITLE OF INVENTION: Regulated Gene Expression in Plants
; FILE REFERENCE: 674538-2001.3
; CURRENT APPLICATION NUMBER: US/10/192.078
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 08/732,348
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: PCT/GB00/02071
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: UK 001580.0
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: UK 001578.4
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: UK 9912635.1
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Consensus zinc finger structure
US-10-192-078-23
```

```
Query Match          77.7%; Score 108; DB 12; Length 29;
Best Local Similarity 76.0%; Pred. No. 1.7e-08;
Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 1 YQCEICGKSFSDKSNLTRHLRIHTG 25
Db 2 YKCEGCKAFSQKSNLTRHQRHTG 26
```

```
RESULT 6
US-10-104-047-2805
; Sequence 2805, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104.047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
```

NUMBER OF SEQ ID NOS: 4096  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO: 2805  
LENGTH: 561  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-104-047-2805

Query Match 76.3%; Score 106; DB 12; Length 561;  
Best Local Similarity 72.0%; Pred. No. 7.8e-07;  
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSFSDKSNLTRHRIHTG 25  
DB 411 YECNEGKSFSEKSTLTGTHRIHTG 435

RESULT 7  
US-10-043-487-390  
Sequence 390, Application US/10043487  
Publication No. US20030055220A1  
GENERAL INFORMATION:  
APPLICANT: HYBRIGENICS  
APPLICANT: Pierre, LEGRAIN  
TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptides  
FILE REFERENCE: B4778A  
CURRENT APPLICATION NUMBER: US/10/043,487  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/261,130  
PRIOR FILING DATE: 2001-01-12  
NUMBER OF SEQ ID NOS: 561  
SOFTWARE: Patent version 3.1  
SEQ ID NO: 390  
LENGTH: 362  
TYPE: PRT  
ORGANISM: Shigella flexneri  
US-10-043-487-390

Query Match 74.8%; Score 104; DB 15; Length 362;  
Best Local Similarity 76.0%; Pred. No. 9.4e-07;  
Matches 19; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSFSDKSNLTRHRIHTG 25  
DB 108 YQCEICGKSFSDKSNLTRHRIHTG 132

RESULT 8  
US-09-864-761-33653  
Sequence 33653, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1  
SEQ ID NO: 33653  
LENGTH: 525  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO Z84476.6  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.8  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EST HUMAN HIT: AF150239.1, EVALUATE 1.00e-106  
OTHER INFORMATION: EST HUMAN HIT: AF150239.1, EVALUATE 1.00e-106  
US-09-864-761-33653

Query Match 74.8%; Score 104; DB 9; Length 525;  
Best Local Similarity 68.0%; Pred. No. 1.4e-06;  
Matches 17; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 YQCEICGKSFSDKSNLTRHRIHTG 25  
DB 246 YECNEGKSFSEKSTLTGTHRIHTG 270

RESULT 9  
US-10-094-749-2365  
Sequence 2365, Application US/10094749  
Publication No. US20030219741A1  
GENERAL INFORMATION:  
APPLICANT: ISOGAI, TAKAO  
APPLICANT: SUGIYAMA, TOMOYASU  
APPLICANT: OTSUKI, TETSUJI  
APPLICANT: WAKAMATSU, AI  
APPLICANT: SATO, HIROYUKI  
APPLICANT: ISHII, SHIZUKO  
APPLICANT: YAMAMOTO, JUN-ICHI  
APPLICANT: ISONO, YUUKO  
APPLICANT: HIO, YURI  
APPLICANT: OTSUKA, KAORU  
APPLICANT: NAGAI, KEIICHI  
APPLICANT: IRIE, RYOTARO



APPLICANT: TAMECHIKA, ICHIRO  
APPLICANT: SEKI, NAOHITO  
APPLICANT: YOSHIKAWA, TSUTOMU  
APPLICANT: OTSUKA, MOTOMYUKI  
APPLICANT: NAGAHARI, KENJI  
APPLICANT: MASUHO, YASUHIRO  
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
FILE REFERENCE: 084335/0160  
CURRENT APPLICATION NUMBER: US/10/094,749  
CURRENT FILING DATE: 2002-03-12  
PRIOR APPLICATION NUMBER: 60/350,435  
PRIOR FILING DATE: 2002-01-24  
PRIOR APPLICATION NUMBER: JP 2001-328381  
PRIOR FILING DATE: 2001-09-14  
NUMBER OF SEQ ID NOS: 3381  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2365  
LENGTH: 576  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-094-749-2365

Query Match 74.8%; Score 104; DB 12; Length 576;  
Best Local Similarity 72.0%; Pred. No. 1.5e-06;  
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSPFSDKSNLTRHLRIHTG 25  
DB 493 YQCECGKAFNQSHTLRHRIHTG 517

RESULT 10  
US-10-104-047-2663  
Sequence 2663, Application US/10104047  
Publication No. US20030236392A1  
GENERAL INFORMATION:  
APPLICANT: HELIX RESEARCH INSTITUTE  
TITLE OF INVENTION: No. US20030236392A1e1 full length CDNA  
FILE REFERENCE: H1-A0105  
CURRENT APPLICATION NUMBER: US/10/104,047  
CURRENT FILING DATE: 2002-03-25  
PRIOR APPLICATION NUMBER:  
PRIOR FILING DATE:  
NUMBER OF SEQ ID NOS: 4096  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2663  
LENGTH: 409  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-104-047-2663

Query Match 74.1%; Score 103; DB 12; Length 409;  
Best Local Similarity 72.0%; Pred. No. 1.5e-06;  
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSPFSDKSNLTRHLRIHTG 25  
DB 357 YQCECGKAFNKSNTLRHLRIHTG 381

RESULT 11  
US-10-104-047-3470  
Sequence 3470, Application US/10104047  
Publication No. US20030236392A1  
GENERAL INFORMATION:  
APPLICANT: HELIX RESEARCH INSTITUTE  
TITLE OF INVENTION: No. US20030236392A1e1 full length CDNA  
FILE REFERENCE: H1-A0105  
CURRENT APPLICATION NUMBER: US/10/104,047  
CURRENT FILING DATE: 2002-03-25  
PRIOR APPLICATION NUMBER:  
PRIOR FILING DATE:  
NUMBER OF SEQ ID NOS: 4096

SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3470  
LENGTH: 475  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-104-047-3470

Query Match 74.1%; Score 103; DB 12; Length 475;  
Best Local Similarity 72.0%; Pred. No. 1.7e-06;  
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSPFSDKSNLTRHLRIHTG 25  
DB 422 YQCECGKAFNKSNTLRHLRIHTG 446

RESULT 12  
US-10-198-677-78  
Sequence 78, Application US/10198677  
Publication No. US20030119023A1  
GENERAL INFORMATION:  
APPLICANT: CHOO, Yen  
APPLICANT: KLUG, Aaron  
APPLICANT: MOORE, Michael  
TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY  
TITLE OF INVENTION: FLEXIBLE LINKERS CONNECTED NUCLEIC ACID BINDING  
FILE REFERENCE: 8325-2011 / G11-US1  
CURRENT APPLICATION NUMBER: US/10/198,677  
CURRENT FILING DATE: 2002-07-17  
NUMBER OF SEQ ID NOS: 144  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 78  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence: consensus structure  
US-10-198-677-78

Query Match 73.4%; Score 102; DB 15; Length 25;  
Best Local Similarity 75.0%; Pred. No. 1.1e-07;  
Matches 18; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSPFSDKSNLTRHLRIHT 24  
DB 2 YQCECGKAFNQSHTLRHRIHT 25

RESULT 13  
US-10-104-047-3434  
Sequence 3434, Application US/10104047  
Publication No. US20030236392A1  
GENERAL INFORMATION:  
APPLICANT: HELIX RESEARCH INSTITUTE  
TITLE OF INVENTION: No. US20030236392A1e1 full length CDNA  
FILE REFERENCE: H1-A0105  
CURRENT APPLICATION NUMBER: US/10/104,047  
CURRENT FILING DATE: 2002-03-25  
PRIOR APPLICATION NUMBER:  
PRIOR FILING DATE:  
NUMBER OF SEQ ID NOS: 4096  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3434  
LENGTH: 530  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-104-047-3434

Query Match 73.4%; Score 102; DB 12; Length 530;  
Best Local Similarity 68.0%; Pred. No. 2.7e-06;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 YOCICGKSFSDKSNLTRHLRIHTG 25  
Db 438 YACTVCGKAFSOGKSNLTREKXHTG 462

## RESULT 14

US-10-029-386-33120  
; Sequence 33120, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: A60MICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; NUMBER OF SEQ ID NOS: 2001-12-20  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 31120  
; LENGTH: 242  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AF205588.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.89  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
; OTHER INFORMATION: SWISSPROT HIT: Q06730, EVALUE 2.00e-84  
US-10-029-386-33120

Query Match 71.9%; Score 100; DB 12; Length 242;  
Best Local Similarity 70.8%; Pred. No. 2.3e-06;  
Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 YOCICGKSFSDKSNLTRHLRIHT 24  
Db 219 YECNCGKSFSEKSTLTKHLRIHT 242

## RESULT 15

US-10-094-749-1859  
; Sequence 1859, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, KYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOTYUKI  
; APPLICANT: NAGAHARA, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381

; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1859  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-749-1859

Query Match 71.9%; Score 100; DB 12; Length 295;  
Best Local Similarity 68.0%; Pred. No. 2.8e-06;  
Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 YOCICGKSFSDKSNLTRHLRIHTG 25  
Db 127 FOCNCGKAFDRSLSLTHRIHTG 151

Search completed: January 6, 2004, 16:17:24  
Job time : 34 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: January 6, 2004, 16:09:06 ; Search time 22 Seconds  
(without alignments)

48.081 Million cell updates/sec

Title: US-09-774-940A-1

Sequence: 1 YOCETGKSPKSKNLTNRHRTHTG 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: Issued Patents AA:\*  
2: /cgn2\_6/prodata/1/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/5A.COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/5B.COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/5A.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	73.4	89	3	US-08-793-408-18
2	102	73.4	89	3	US-09-139-762A-18
3	96	69.1	26	2	US-08-620-151-84
4	94	67.6	671	3	US-09-121-321-16
5	94	67.6	671	3	US-08-933-803A-16
6	93	66.9	161	4	US-09-389-956-64
7	93	66.9	796	4	US-09-389-956-64
8	92	66.2	26	2	US-08-491-527A-15
9	92	66.2	26	2	US-08-620-151-57
10	92	66.2	26	2	US-08-620-151-58
11	92	66.2	26	2	US-08-620-151-89
12	92	66.2	41	4	US-09-389-831-15
13	92	66.2	488	3	US-08-933-750C-17
14	92	66.2	488	3	US-09-234-613-17
15	91	65.5	56	4	US-08-733-622C-33
16	91	65.5	208	4	US-08-733-622C-37
17	91	65.5	209	4	US-08-733-622C-8
18	91	65.5	507	4	US-08-733-622C-2
19	91	65.5	640	3	US-09-262-773-4
20	91	65.5	648	3	US-09-262-773-2
21	90	64.7	26	2	US-08-620-151-78
22	89	64.0	25	2	US-08-620-151-98
23	89	64.0	543	4	US-09-362-123A-4
24	88	63.3	26	2	US-08-620-151-30
25	88	63.3	26	2	US-08-620-151-37
26	88	63.3	711	2	US-08-820-170A-10
27	88	63.3	711	3	US-09-055-699-10

28	88	63.3	711	3	US-09-273-565-10	Sequence 10, Appl
29	88	63.3	711	4	US-09-565-538-10	Sequence 10, Appl
30	88	63.3	711	4	US-09-661-468-10	Sequence 10, Appl
31	88	63.3	711	4	US-09-976-165-10	Sequence 10, Appl
32	87.5	62.9	33	3	US-08-793-408-12	Sequence 12, Appl
33	87.5	62.9	33	3	US-09-139-762A-12	Sequence 12, Appl
34	87	62.6	56	3	US-08-620-151-53	Sequence 187, App
35	87	62.6	56	3	US-08-711-417C-187	Sequence 189, App
36	87	62.6	56	3	US-08-711-417C-188	Sequence 189, App
37	87	62.6	56	3	US-08-711-417C-189	Sequence 189, App
38	87	62.6	60	2	US-08-117-952-786	Sequence 201, App
39	87	62.6	334	4	US-08-711-417C-201	Sequence 201, App
40	87	62.6	334	4	US-08-733-622C-23	Sequence 23, Appl
41	87	62.6	338	3	US-08-933-750C-4	Sequence 4, Appl
42	87	62.6	338	3	US-09-234-613-4	Sequence 196, App
43	87	62.6	461	3	US-08-711-417C-196	Sequence 25, Appl
44	87	62.6	461	4	US-08-733-622C-25	Sequence 117, App
45	87	62.6	462	3	US-08-486-099-117	

#### ALIGNMENTS

RESULT 1  
US-08-793-408-18  
Sequence 18, Application US/08793408  
Patent No. 6007988  
GENERAL INFORMATION:  
APPLICANT: Choo, Yen  
APPLICANT: Kluug, Aaron  
APPLICANT: Sanchez Garcia, Isidro  
TITLE OF INVENTION: Improvements in or Relating to  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Pillsbury Madison & Sutro, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/793,408  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/01949  
FILING DATE: 17-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9514698.1  
FILING DATE: 18-JUL-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9422534.9  
FILING DATE: 08-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9416880.4  
FILING DATE: 20-AUG-1994  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 89 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-793-408-18  
Query Match 73.4%; Score 102; DB 3; Length 89;  
Best Local Similarity 68.0%; Pred. No. 7.3e-08;

Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDSKSNLTRHRIHTG 25  
Db 63 FQCRICWPNFSDRSNLTFRHRIHTG 87

## RESULT 2

US-09-139-762A-18  
Sequence 18, Application US/09139762A  
Patent No. 6013453  
GENERAL INFORMATION:  
APPLICANT: Choo, Yen  
APPLICANT: Kling, Aaron  
APPLICANT: Sanchez Garcia, Isidro  
TITLE OF INVENTION: Improvements in or Relating to  
TITLE OF INVENTION: Binding Proteins for Recognition of DNA  
NUMBER OF SEQUENCES: 125  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/139, 762A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/793,408  
FILING DATE: 02-JUN-1997  
APPLICATION NUMBER: PCT/GB95/01949  
FILING DATE: 17-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9514698.1  
FILING DATE: 18-JUL-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9422534.9  
FILING DATE: 08-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9416880.4  
FILING DATE: 20-AUG-1994  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 89 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-139-762A-18

Query Match 73.4%; Score 102; DB 3; Length 89;  
Best Local Similarity 68.0%; Pred. No. 7.3e-08;  
Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDSKSNLTRHRIHTG 25  
Db 63 FQCRICWPNFSDRSNLTFRHRIHTG 87

## RESULT 3

US-08-620-151-84  
Sequence 84, Application US/08620151  
Patent No. 5928855  
GENERAL INFORMATION:  
APPLICANT: Imperiali, Barbara  
APPLICANT: Walkup, Grant K.

TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR  
TITLE OF INVENTION: DIVALENT ZINC  
NUMBER OF SEQUENCES: 136  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE  
STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60611-5599

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent'n Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/620,151  
FILING DATE: 22-MAR-1996  
CLASSIFICATION: 422  
ATTORNEY/AGENT INFORMATION:  
NAME: Shannon, Karen L.  
REGISTRATION NUMBER: 36,675  
REFERENCE/DOCKET NUMBER: 8597/6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-321-4200  
TELEFAX: 312-321-4299  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-620-151-84

Query Match 69.1%; Score 96; DB 2; Length 26;  
Best Local Similarity 60.0%; Pred. No. 1.5e-07;  
Matches 15; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 YQCEICGKSPDSKSNLTRHRIHTG 25  
Db 2 YECVQCKTFSHKANLIKQRIHTG 26

## RESULT 4

US-09-121-321-16  
Sequence 16, Application US/09121321  
Patent No. 6090783  
GENERAL INFORMATION:  
APPLICANT: Saiga, Akiniko  
APPLICANT: Oyita, Satoshi  
APPLICANT: Igarashi, Hisanaga  
APPLICANT: Okumura, Kouichi  
APPLICANT: Sakaguchi, Gaku  
TITLE OF INVENTION: DNA MOLECULE RELATING TO SUPPRESSION OF  
TITLE OF INVENTION: GENE EXPRESSION AND NOVEL PROTEIN  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent'n Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/121,321  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/933,803  
FILING DATE: 19-SEP-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley, James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: SHGN-12CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 671 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-121-321-16

Query Match 67.6%; Score 94; DB 3; Length 671;  
Best Local Similarity 64.0%; Pred. No. 9.3e-06;  
Matches 16; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 YQCEICGKSPSDKSNLTRRLRIHTG 25  
Db 602 FQCALGKSFIRKQNLKHKRIHTG 626

RESULT 5  
US-08-933-803A-16  
Sequence 16, Application US/08933803A  
Patent No. 6218522  
GENERAL INFORMATION:  
APPLICANT: Saiga, Akiniko  
APPLICANT: Orita, Satoshi  
APPLICANT: Igarashi, Hisanaga  
APPLICANT: Okumura, Kouichi  
APPLICANT: Sakaguchi, Gaku  
TITLE OF INVENTION: DNA MOLECULE RELATING TO SUPPRESSION OF  
TITLE OF INVENTION: GENE EXPRESSION AND NOVEL PROTEIN  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,803A  
FILING DATE: 19-SEP-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley, James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: SHGN-12CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 671 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-933-803A-16

Query Match 67.6%; Score 94; DB 3; Length 671;

Best Local Similarity 64.0%; Pred. No. 9.3e-06;  
Matches 16; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 YQCEICGKSPSDKSNLTRRLRIHTG 25  
Db 602 FQCALGKSFIRKQNLKHKRIHTG 626

RESULT 6  
US-09-389-956-64  
Sequence 64, Application US/09389956  
Patent No. 6586579  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: PR-Domain Containing Nucleic Acids, Polypeptides,  
TITLE OF INVENTION: Antibodies and Methods  
FILE REFERENCE: P-17 3611  
CURRENT APPLICATION NUMBER: US/09/389,956  
CURRENT FILING DATE: 1999-09-03  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 64  
LENGTH: 161  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-389-956-64

Query Match 66.9%; Score 93; DB 4; Length 161;  
Best Local Similarity 60.0%; Pred. No. 2.8e-06;  
Matches 15; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSPSDKSNLTRRLRIHTG 25  
Db 27 HKCDFSKAFSDPSNLRTHLRIHTG 51

RESULT 7  
US-09-389-956-2  
Sequence 2, Application US/09389956  
Patent No. 6586579  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: PR-Domain Containing Nucleic Acids, Polypeptides,  
TITLE OF INVENTION: Antibodies and Methods  
FILE REFERENCE: P-17 3611  
CURRENT APPLICATION NUMBER: US/09/389,956  
CURRENT FILING DATE: 1999-09-03  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 2  
LENGTH: 796  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-389-956-2

Query Match 66.9%; Score 93; DB 4; Length 796;  
Best Local Similarity 60.0%; Pred. No. 1.6e-05;  
Matches 15; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 YQCEICGKSPSDKSNLTRRLRIHTG 25  
Db 613 HKCDFSKAFSDPSNLRTHLRIHTG 637

RESULT 8  
US-08-491-527A-15  
Sequence 15, Application US/08491527A  
Patent No. 5824483  
GENERAL INFORMATION:  
APPLICANT: Houston, Michael E.  
APPLICANT: Hodges, Robert S.  
TITLE OF INVENTION: Conformationally-Restricted Combinatorial  
TITLE OF INVENTION: Library Composition and Method

NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/491,527A  
FILING DATE: 16-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,199  
FILING DATE: 15-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/245,507  
FILING DATE: 18-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 7900-0008.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: peptide CPL  
US-08-491-527A-15

Query Match 66.2%; Score 92; DB 2; Length 26;  
Best Local Similarity 64.0%; Pred. No. 5.6e-07;  
Matches 16; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YQCEIGKSPDKNLTHRIHTG 25  
Db 2 YKCEGKSPQSKSDLVKHORHTG 26

RESULT 9  
US-08-620-151-57  
Sequence 57, Application US/08620151  
Patent No. 5928955  
GENERAL INFORMATION:  
APPLICANT: Imperiali, Barbara  
APPLICANT: Walkup, Grant K.  
TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR  
DIVALENT ZINC  
NUMBER OF SEQUENCES: 136  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE  
STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60611-5599  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/620,151  
FILING DATE: 22-MAR-1996  
CLASSIFICATION: 422  
ATTORNEY/AGENT INFORMATION:  
NAME: Shannon, Karen L.  
REGISTRATION NUMBER: 36,675  
REFERENCE/DOCKET NUMBER: 8597/6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-321-4200  
TELEFAX: 312-321-4299  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-620-151-57

Query Match 66.2%; Score 92; DB 2; Length 26;  
Best Local Similarity 64.0%; Pred. No. 5.6e-07;  
Matches 16; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YQCEIGKSPDKNLTHRIHTG 25  
Db 2 YKCEGKSPQSKSDLVKHORHTG 26

RESULT 10  
US-08-620-151-58  
Sequence 58, Application US/08620151  
Patent No. 5928955  
GENERAL INFORMATION:  
APPLICANT: Imperiali, Barbara  
APPLICANT: Walkup, Grant K.  
TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR  
DIVALENT ZINC  
NUMBER OF SEQUENCES: 136  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE  
STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60611-5599  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/620,151  
FILING DATE: 22-MAR-1996  
CLASSIFICATION: 422  
ATTORNEY/AGENT INFORMATION:  
NAME: Shannon, Karen L.  
REGISTRATION NUMBER: 36,675  
REFERENCE/DOCKET NUMBER: 8597/6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-321-4200  
TELEFAX: 312-321-4299  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-620-151-58

Query Match 66.2%; Score 92; DB 2; Length 26;  
Best Local Similarity 64.0%; Pred. No. 5.6e-07;  
Matches 16; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 YQCEICGKSPSDKSNLTRHLRIHTG 25  
Db 2 YACSTCGKSTFKSDLAHQRIHTG 26

RESULT 11  
US-08-620-151-89  
Sequence 89, Application US/08620151  
Patent No. 5928855

GENERAL INFORMATION:  
APPLICANT: Imperiali, Barbara  
APPLICANT: Walkup, Grant K.  
TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR  
TITLE OF INVENTION: DIVALENT ZINC  
NUMBER OF SEQUENCES: 136  
CORRESPONDENCE ADDRESS:

ADDRESSEE: BRINKS, HOFER, GILSON & LIONE  
STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
STREET: Plaza Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60611-5599

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/620,151  
FILING DATE: 22-MAR-1996  
CLASSIFICATION: 422

ATTORNEY/AGENT INFORMATION:

NAME: Shannon, Karen L.  
REGISTRATION NUMBER: 36,675  
REFERENCE/DOCKET NUMBER: 8597/6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-321-4200  
TELEFAX: 312-321-4299

INFORMATION FOR SEQ ID NO: 89:

SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-620-151-89

Query Match 66.2%; Score 92; DB 2; Length 26;  
Best Local Similarity 64.0%; Pred. No. 5.6e-07;  
Matches 16; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YQCEICGKSPSDKSNLTRHLRIHTG 25  
Db 2 YECTCGKTFQSRSTLRLHRIHTG 26

RESULT 12  
US-09-389-831-15  
Sequence 15, Application US/09389831  
Patent No. 6521420

GENERAL INFORMATION:

APPLICANT: Herman, Ira M.

APPLICANT: Slesciakiewicz, Greg J.

TITLE OF INVENTION: HYPEREXTENSION ASSOCIATED TRANSCRIPTION FACTORS AND USES THEREFOR

FILE REFERENCE: MBI-006

CURRENT APPLICATION NUMBER: US/09/389,831

CURRENT FILING DATE: 1999-09-03

NUMBER OF SEQ ID NOS: 19

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 15

LENGTH: 41

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Description of Artificial Sequence: synthetic

OTHER INFORMATION: primer

US-09-389-831-15

Query Match 66.2%; Score 92; DB 4; Length 41;  
Best Local Similarity 60.0%; Pred. No. 9.2e-07;  
Matches 15; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YQCEICGKSPSDKSNLTRHLRIHTG 25  
Db 4 YQCSLCKAFQSRSSILVGHRIHTG 28

RESULT 13  
US-08-933-750C-17  
Sequence 17, Application US/08933750C  
Patent No. 5932442

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Hillman, Jennifer L.

APPLICANT: Bandman, Olga

APPLICANT: Shah, Purni

APPLICANT: Au-Young, Janice

APPLICANT: Yue, Henry

APPLICANT: Guebler, Karl J.

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: HUMAN REGULATORY MOLECULES

NUMBER OF SEQUENCES: 98

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/933,750C

FILING DATE: September 23, 1997

CLASSIFICATION: 536

PRIOR APPLICATION NUMBER:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0356 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELFX:

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 488 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: HBSTNOT03

CLONE: 641127

US-08-933-750C-17

Query Match 66.2%; Score 92; DB 2; Length 488;